

```
Qy 2218 TAATTTTATGCTGAAACAGAGTTTGGCGACATTTGGACCATCAGAAAGCAGAAAGTGTAC 2277
Db 783 TAATTTTATGCTGAAACAGAGTTTGGCGACATTTGGACCATCAGAAAGCAGAAAGTGTAC 842
Qy 2278 TAATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGTATGCGCAG 2337
Db 843 TAATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGTATGCGCAG 902
Qy 2338 TGATCCGAGTACTTTTGGAGCCCAAGACAGGTGTGATCTCTTGGAGCCAGGAGTTTGGAGC 2397
Db 903 TGATCCGAGTACTTTTGGAGCCCAAGACAGGTGTGATCTCTTGGAGCCAGGAGTTTGGAGC 962
Qy 2398 CAGACTGCACAAACAGCAGTGTGAGTGTCTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGTATGCGCAG 2457
Db 963 CAGACTGCACAAACAGCAGTGTGAGTGTCTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGTATGCGCAG 1022
Qy 2458 GGTGGTGACACCTGTGTAGTCCAGGTACTCAGGAGGCTGAGG-TAGTAGGATTTGTTGAG 2516
Db 1023 GGTGGTGACACCTGTGTAGTCCAGGTACTCAGGAGGCTGAGGTTATTTGGGATTTGTTGAG 1082
Qy 2517 ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 2560
Db 1083 ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 1126

RESULT 15
CO494252
LOCUS CO494252 1128 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 26119 from Patent WO0160860.
ACCESSION CO494252
VERSION CO494252.1 GI:41459871
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Schlegel, R., Endege, W.O. and Monahan, J.E.
Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 26119 23-AUG-2001;
AUTHORS Millennium Predictive Medicine, Inc. (US)
TITLES Location/Qualifiers
FEATURES
source
1..1128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 28.4%; Score 1068; DB 6; Length 1128;
Best Local Similarity 98.0%; Pred. No. 1.2e-176;
Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;
Qy 1438 CCAATTTCAATGCGCAACATCTCCCATCCCAATACCTATTTTCTTTGAGGTGAAT 1497
Db 4 CTACAAATTTAAATGCAACATCTGCCATCCCAA-TACCTATTTTCATTTTGGGGGTAA 62
Qy 1498 CTACTCAATAAATGTTGAAGACCACTGACAGACCCCTTTGCTAACCTGACATTTACTTC 1557
Db 63 CTACTCAATAAATGTTGAAGACCACTGACAGACCCCTTTGCTAACCTGACATTTACTTC 122
Qy 1558 AATTTTCTTTTCTATGTAAGTATTTTGGATATTTTGGATATAAATCTTGCAGTAATAGTTCAA 1617
Db 123 AATTTTCTTTTCTATGTAAGTATTTTGGATATTTTGGATATAAATCTTGCAGTAATAGTTCAA 182
Qy 1618 AATTAATAGTTTTCATGCTGCTTTCTGAGAGAGAAATTTGAAGTGTCAAAATTA 1677
Db 183 AATTAATAGTTTTCATGCTGCTTTCTGAGAGAGAAATTTGAAGTGTCAAAATTA 242
Qy 1678 AAAAGATGAATGAAGCATATATAATTTCTCAATTTTTTCAATTTTTTCTAGTCAACAGAGA 1737
Db 243 AAAAGATGAATGAAGCATATATAATTTCTCAATTTTTTCAATTTTTTCTAGTCAACAGAGA 302
```

```
Qy 1738 ATCGAAGGATCTGTTCAAATATATAGTAATAAATTTGAAAAATTAACCTTGTGCTTATATTTTG 1797
Db 303 ATCGAAGGATCTGTTCAAATATATAGTAATAAATTTGAAAAATTAACCTTGTGCTTATATTTTG 362
Qy 1798 TTTTGCACACACACTAGTTTAAATTTTAACTGTGACCTAGTTTATCTTACCGAAGGTGTGATGTGT 1857
Db 363 TTTTGCACACACACTAGTTTAAATTTTAACTGTGACCTAGTTTATCTTACCGAAGGTGTGATGTGT 422
Qy 1858 AGTTTCTGTTTAAAAATTTCAAGCAAACTGCGAAAAATTAATCCATCTAATTTATGCTTTCTTT 1917
Db 423 AGTTTCTGTTTAAAAATTTCAAGCAAACTGCGAAAAATTAATCCATCTAATTTATGCTTTCTTT 482
Qy 1918 CCCAAGAGTTTAAATATATATGTCAGCTTCCCTAAATTTGGAGACAAAGCCTTAATTTGA 1977
Db 483 CCCAAGAGTTTAAATATATATGTCAGCTTCCCTAAATTTGGAGACAAAGCCTTAATTTGA 542
Qy 1978 CAATGCATTCATATATATATTTTGTATAGTTTACAGTATACAGAGTTTGAATATCCCTTAG 2037
Db 543 CAATGCATTCATATATATATTTTGTATAGTTTACAGTATACAGAGTTTGAATATCCCTTAG 602
Qy 2038 ATGAGATGCTTGGGACCCAGAGTGTTTTGGATTTTCAAGTTTATTTTGGATTTTGGAAATA 2097
Db 603 ATGAGATGCTTGGGACCCAGAGTGTTTTGGATTTTCAAGTTTATTTTGGATTTTGGAAATA 662
Qy 2098 TTTCCATACATATATATGAGAGAGTTTGGAAAAATGGGATTCAGTCTAATCAATAAATTCAC 2157
Db 663 TTTCCATACATATATATGAGAGAGTTTGGAAAAATGGGATTCAGTCTAATCAATAAATTCAC 722
Qy 2158 TTATGTTTGTATATACACCTTATCTGAATAGCTTCAAGGTAAATTTTATACAATAATTTTAAA 2217
Db 723 TTATGTTTGTATATACACCTTATCTGAATAGCTTCAAGGTAAATTTTATACAATAATTTTAAA 782
Qy 2218 TAAATTTTATGCTGAAACAGAGTTTGGCGCAATTTGGACCATTCAGAAAGCAGAAAGTGTAC 2277
Db 783 TAAATTTTATGCTGAAACAGAGTTTGGCGCAATTTGGACCATTCAGAAAGCAGAAAGTGTAC 842
Qy 2278 TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTTAACTGGTGTATGATGCTTATGCCAG 2337
Db 843 TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTTAACTGGTGTATGATGCTTATGCCAG 902
Qy 2338 TGATCCGAGTACTTTTGGAGCCCAAGACAGGTGTGATCTCTTGGAGCCAGGAGTTTGGAGC 2397
Db 903 TGATCCGAGTACTTTTGGAGCCCAAGACAGGTGTGATCTCTTGGAGCCAGGAGTTTGGAGC 962
Qy 2398 CAGACTGCACAAACAGCAGTGTGAGACCTCGTTTCTCAAAAATTTAAAAATTTAGCCAGGTGT 2457
Db 963 CAGACTGCACAAACAGCAGTGTGAGACCTCGTTTCTCAAAAATTTAAAAATTTAGCCAGGTGT 1022
Qy 2458 GGTGGTGACACCTGTGTAGTCCAGGTACTCAGGAGGCTGAGG-TAGTAGGATTTGTTGAG 2516
Db 1023 GGTGGTGACACCTGTGTAGTCCAGGTACTCAGGAGGCTGAGGTATTGGGATTTGTTGAG 1082
Qy 2517 ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 2560
Db 1083 ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 1126
```

Search completed: November 18, 2004, 18:25:07  
Job time : 15382 secs

Db 123 AATTCTCTCTTTCTATGTACTGGATATATTTTGCATATATAAACTTGCAGTAATAGTTCAAA 182  
 QY 1618 AATTATAGTTTTTGCATTTGGCTTTTCTGAGAAGAAATTTGAAAGTGTACAAAATAA 1677  
 Db 183 AATTAAATAGTTTTTGCATTTGGCTTTTCTGAGAAGAAATTTGAAAGTGTACAAAATAA 242  
 QY 1678 AAAAAGATGAATGAAGCATATATAAATTTGTCATATTTTCAATTTTCTAGTCAACAGAGA 1737  
 Db 243 AAAAAGATGAATGAAGCATATATAAATTTGTCATATTTTCAATTTTCTAGTCAACAGAGA 302  
 QY 1738 ATCGAAGGATCTGTTCAATATATAGTAAATAATTTGAAATTTGAAATTTGAAATTTG 1797  
 Db 303 ATCGAAGGATCTGTTCAATATATAGTAAATAATTTGAAATTTGAAATTTGAAATTTG 362  
 QY 1798 TTGCAACACACACTAGTTAAATTTTAACTGCTGACTAGTTATCTCTACCGAAGTGTATGTT 1857  
 Db 363 TTGCAACACACACTAGTTAAATTTTAACTGCTGACTAGTTATCTCTACCGAAGTGTATGTT 422  
 QY 1858 AGTTCTCTGTTTTAAATTTCAAGCAAACTGGAAATAATCCATCTAATTTATGCTTTCTTT 1917  
 Db 423 AGTTCTCTGTTTTAAATTTCAAGCAAACTGGAAATAATCCATCTAATTTATGCTTTCTTT 482  
 QY 1918 CCAAGAAGTTTTTAAATGATATGCGAGTTCTTAAATTTGGAGACAAAAGCCTTAATGA 1977  
 Db 483 CCAAGAAGTTTTTAAATGATATGCGAGTTCTTAAATTTGGAGACAAAAGCCTTAATGA 542  
 QY 1978 CAATGCATTCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 2037  
 Db 543 CAATGCATTCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 602  
 QY 2038 ATGAGATGCTTGGGACCAAGAGTTTGGATTTTCAATTTTGGATTTTGGAAATA 2097  
 Db 603 ATGAGATGCTTGGGACCAAGAGTTTGGATTTTCAATTTTGGATTTTGGAAATA 662  
 QY 2098 TTTCCATACATATAATGAGAGAGTTGGTAAATGGGATTCAGTCTAATCAATAAATTCAC 2157  
 Db 663 TTTCCATACATATAATGAGAGAGTTGGTAAATGGGATTCAGTCTAATCAATAAATTCAC 722  
 QY 2158 TTATGTTTGATATACACCTTTATCTGAATAGCTTGAAGTAAATTTTATACATATTTTAA 2217  
 Db 723 TTATGTTTGATATACACCTTTATCTGAATAGCTTGAAGTAAATTTTATACATATTTTAA 782  
 QY 2218 TAATTTTATGCTGAAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAAAGTGTAC 2277  
 Db 783 TAATTTTATGCTGAAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAAAGTGTAC 842  
 QY 2278 TATTTCAAGTCAGTGTCTAAAAGTTTCAAGTGTAAAGTGTGATGTCAGTTCATGCCAG 2337  
 Db 843 TATTTCAAGTCAGTGTCTAAAAGTTTCAAGTGTAAAGTGTGATGTCAGTTCATGCCAG 902  
 QY 2338 TGATCCGAGTACTTTGGGAAGCCAGACAGAGTGTCTCTTGAAGCCAGAGTTTGGAGC 2397  
 Db 903 TGATCCGAGTACTTTGGGAAGCCAGACAGAGTGTCTCTTGAAGCCAGAGTTTGGAGC 962  
 QY 2398 CAGACTGCAACACAGAGTGTCTCTTCTCAATAATTTAAATTTAGCCAGGTTGT 2457  
 Db 963 CAGACTGCAACACAGAGTGTCTCTTCTCAATAATTTAAATTTAGCCAGGTTGT 1022  
 QY 2458 GGTGGTGCACACTGTAGTCCAGGTTACTCAGAGAGCTGAGG-TAGTAGATTTGTTGAG 2516  
 Db 1023 GGTGGTGCACACTGTAGTCCAGGTTACTCAGAGAGCTGAGGTTATTTGGGATTTGTTGAG 1082  
 QY 2517 ACTGGGAGGTTGAGGTTGAACTGAGCAGAGTCTTGCACCAACA 2560  
 Db 1083 ACTGGGAGGTTGAGGTTGAACTGAGCAGAGTCTTGCACCAACA 1126

RESULT 14  
 CQ488930  
 LOCUS 1128 bp DNA linear PAT 30-JAN-2004  
 DEFINITION Sequence 20797 from Patent WO0160860.  
 CQ488930  
 ACCESSION CQ488930.1 GI:41454549  
 VERSION

KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 Schlegel R., Endege, W.O. and Monahan, J.E.  
 AUTHORS Genes differentially expressed in human prostate cancer and their use  
 TITLE Patent: WO 0160860-A 20797 23-AUG-2001;  
 JOURNAL Millennium Predictive Medicine, Inc. (US)  
 FEATURES Location/Qualifiers  
 source 1..1128  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 28.4%; Score 1068; DB 6; Length 1128;  
 Best Local Similarity 98.0%; Pred. No. 1.2e-176;  
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;  
 QY 1438 CCAAAATTTCAAATGCCAACATCTCCCATCCAAATTTACCTATTTTCACTTTGAGGTGAAT 1497  
 Db 4 CTACAAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTCACTTTGAGGGTGAAT 62  
 QY 1498 CTACTCAATAACTGTGTAAAGACAGTGCAGACACCTTTGCTAACTGCATTTACTTC 1557  
 Db 63 CTACTCAATAACTGTGTAAAGACAGTGCAGACACCTTTGCTAACTGCATTTACTTC 122  
 QY 1558 AATTTTCTTTTCTATGTACTGATATATTTTGCATATAAATTTGAGTAAATAGTTCAAA 1617  
 Db 123 AATTTTCTTTTCTATGTACTGATATATTTTGCATATAAATTTGAGTAAATAGTTCAAA 182  
 QY 1618 AATTAATAGTTTTTGACATTTGGCTTTTCTGAGAAGAAATTTGAAAGTGTACAAAATAA 1677  
 Db 183 AATTAATAGTTTTTGACATTTGGCTTTTCTGAGAAGAAATTTGAAAGTGTACAAAATAA 242  
 QY 1678 AAAAAGATGAATGAAGCATATATAAATTTGTCATATTTTCAATTTTCTAGTCAACAGAGA 1737  
 Db 243 AAAAAGATGAATGAAGCATATATAAATTTGTCATATTTTCAATTTTCTAGCACAACAGAGA 302  
 QY 1738 ATCGAAGGATCTGTTCAATATATAGTAAATAATTTGAAATTTGAAATTTGAAATTTG 1797  
 Db 303 ATCGAAGGATCTGTTCAATATATAGTAAATAATTTGAAATTTGAAATTTGAAATTTG 362  
 QY 1798 TTTGCAACACACTAGTTAAATTTTAACTGCTGACTAGTTATCTCTACCGAAGTGTATGTT 1857  
 Db 363 TTTGCAACACACTAGTTAAATTTTAACTGCTGACTAGTTATCTCTACCGAAGTGTATGTT 422  
 QY 1858 AGTTCTCTGTTTTAAATTTCAAGCAAACTGGAAATAATCCATCTAATTTATGCTTTCTTT 1917  
 Db 423 AGTTCTCTGTTTTAAATTTCAAGCAAACTGGAAATAATCCATCTAATTTATGCTTTCTTT 482  
 QY 1918 CCAAGAAGTTTTTAAATGATATGCGAGTTCTTAAATTTGGAGACAAAAGCCTTAATGA 1977  
 Db 483 CCAAGAAGTTTTTAAATGATATGCGAGTTCTTAAATTTGGAGACAAAAGCCTTAATGA 542  
 QY 1978 CAATGCATTCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 2037  
 Db 543 CAATGCATTCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 602  
 QY 2038 ATGAGATGCTTGGGACCAAGAGTTTGGATTTTCAATTTTGGATTTTGGAAATA 2097  
 Db 603 ATGAGATGCTTGGGACCAAGAGTTTGGATTTTCAATTTTGGATTTTGGAAATA 662  
 QY 2098 TTTCCATACATATAATGAGAGAGTTGGTAAATGGGATTCAGTCTAATCAATAAATTCAC 2157  
 Db 663 TTTCCATACATATAATGAGAGAGTTGGTAAATGGGATTCAGTCTAATCAATAAATTCAC 722  
 QY 2158 TTATGTTTGATATACACCTTTATCTGAATAGCTTGAAGTAAATTTTATACATATTTTAA 2217  
 Db 723 TTATGTTTGATATACACCTTTATCTGAATAGCTTGAAGTAAATTTTATACATATTTTAA 782

```
Db 1083 ACTGGAGTTGAGGCTGAAGTGAAGCCAGGATCTTGGCCACACA 1126

RESULT 12
CQ488563
LOCUS CQ488563 1128 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 20430 from Patent WO0160860.
ACCESSION CQ488563
VERSION CQ488563.1 GI:41454182
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 20430 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
source Location/Qualifiers
1..1128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 28.4%; Score 1068; DB 6; Length 1128;
Best Local Similarity 98.0%; Pred. No. 1.2e-176;
Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

Qy 1438 CCAATTTCAAATGCGACATCTCCCATCCAATTACCTTTGAGGTGTAAT 1497
Db 4 CTACAAATTTAAATGCAACATCTGCCCATCAA-TACCTATTTCATCTTTGAGGGTAAT 62

Qy 1498 CTACTCAATAAATGTAAGCAGTGCACAGACCCCTTCTGCTAACCTGCATTTACTTTC 1557
Db 63 CTACTCAATAAATGTAAGCAGTGCACAGACCCCTTCTGCTAACCTGCATTTACTTTC 122

Qy 1558 AATTTTCTTTTCTATGTAAGTATTTTGCATATAAACTTTCAGTAATAGTTTCAAA 1617
Db 123 AATTTTCTTTTCTATGTAAGTATTTTGCATATAAACTTTCAGTAATAGTTTCAAA 182

Qy 1618 AATTAATAGTTTTCAGATTCGCTTTCTGAGAGAGAAATGAAAGTGCACAAAATAA 1677
Db 183 AATTAATAGTTTTCAGATTCGCTTTCTGAGAGAGAAATGAAAGTGCACAAAATAA 242

Qy 1678 AAAAGATGAATGAAGCATATATAATTTGTCAAATTTTCAATTTTCTAGTCAACAGAGA 1737
Db 243 AAAAGATGAATGAAGCATATATAATTTGTCAAATTTTCTAGTCAACAGAGA 302

Qy 1738 ATCGAAGATTCTGTTCAAAATATTAGTAAATAATGAAATAAACTTGTGCTTATATTG 1797
Db 303 ATCGAAGATTCTGTTCAAAATATTAGTAAATAATGAAATAAACTTGTGCTTATATTG 362

Qy 1798 TTTCGACACACTAGTTAATTTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTT 1857
Db 363 TTTCGACACACTAGTTAATTTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTT 422

Qy 1858 AGTTTCTGGTTTAAATTTCAAGCAAACTGGAAATAATCCATCTAATTTATGCTTTCTTT 1917
Db 423 AGTTTCTGGTTTAAATTTCAAGCAAACTGGAAATAATCCATCTAATTTATGCTTTCTTT 482

Qy 1918 CCCAAGAAGTTTTPAATGATATGATGCGACCTTCTTAATTTGGAGACAAAGCCTTAAATGA 1977
Db 483 CCCAAGAAGTTTTPAATGATATGATGCGACCTTCTTAATTTGGAGACAAAGCCTTAAATGA 542

Qy 1978 CAATGCATTCAATATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 2037
Db 543 CAATGCATTCAATATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 602

Qy 2038 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTCAGATTTATTTTGGATTTTGGATA 2097
Db 1126
```

Db	1301	AGCTGAACAAATGTCACATTCATATATTGATTTAATGCTTGTGTTAACTCCTGTCT	1360
Qy	1237	CCCTAAATGATATAAAGTAACGGTGACCTGACCACTCAGGACACATTTCTCAGGACC	1296
Db	1361	CCCTAAGATGATATAAAGTAACGGTGACCTGACCACTCAGGACACATTTCTCAGGACC	1420
Qy	1297	TCTGAGAGTGATCCAGGCCATGTAAGTCAATGTTGGCTCAGATCAACCTCTTTAA	1356
Db	1421	TCTGAGAGTGATCCAGGCCATGTAAGTCAATGTTGGCTCAGATCAACCTCTTTAA	1480
Qy	1357	TATTTTACAGAAATTTGGCTTTTGGTTACCAATAAGTCTCCACAAATATATGTCACAGAA	1416
Db	1481	TATTTTACAGAAATTTGGCTTTTGGTTACCAATAAGTCTCCACAAATATATGTCACAGAA	1540
Qy	1417	CTTCAATTCAGCCCTGCTCACCAAAATTTCAAAATGCGCAACATCTCCCATCCAAATACCT	1476
Db	1541	CTTCAATTCAGCCCTGCTCACCAAAATTTCAAAATGCGCAACATCTCCCATCCAAATACCT	1600
Qy	1477	ATTTCATCTTTGAGGTGTAATCTACTCAATAAAGTGTGAAGACCACTGACAGACCCCT	1536
Db	1601	ATTTCATCTTTGAGGTGTAATCTACTCAATAAAGTGTGAAGACCACTGACAGACCCCT	1660
Qy	1537	TGCTAACTGACATTTACTTCAATTTTCTTTTCTATGTAAGTATTTTTCATATA	1596
Db	1661	TGCTAACTGATATTTACTTCAATTTTCTTTTCTATGTAAGTATTTTTCATATA	1720
Qy	1597	AACCTGAGTAATAGTTCACAAATTAATAGTCTTTTGCATGCTGCTTCTGAGAGAGAA	1656
Db	1721	AACCTGAGTAATAGTTCACAAATTAATAGTCTTTTGCATGCTGCTTCTGAGAGAGAA	1779
Qy	1657	ATTGAAGGTGTCACAAAT---AAAAAAGATGAAATGAGCATATATATTTGCAATTT	1713
Db	1780	ATTGAAGGTGTCACAAATTAATAGTCTTTTGCATGCTGCTTCTGAGAGAGAA	1839
Qy	1714	TTTCAATTTTCTAGTCAACAGAGAAATCGAAGGATCTGTTTCAAAATATATAGTAAAAATGA	1773
Db	1840	AA	1899
Qy	1774	AAATAAA 1780	
Db	1900	AAAAAAA 1906	
RESULT 11			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1102; Conservative			
0; Mismatches			
20; Indels			
2; Gaps			
2; Length 1128;			
28.4%; Score 1068; DB 6;			
Pred. No. 1.2e-176;			
1438 CCAAAATTCAAATGCCAACATCTCCCATCCAAATACCTATTATTCATCTTTGAGGTGAAT			
Qy	1438	CCAAATTCAAATGCCAACATCTCCCATCCAAATACCTATTATTCATCTTTGAGGTGAAT	1497

Db	4	CTACAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTTCATCTTTGAGGGTAAT	62
Qy	1498	CTACTCAATAAACATGTTGTAAGACCACTGACCACTTGTCTAACTGACATTTACTTC	1557
Db	63	CTACTCAATAAACATGTTGTAAGACCACTGACCACTTGTCTAACTGACATTTACTTC	132
Qy	1558	AATTTTCTCTTTCTATGTAAGTATTTTGCATATAAACTTGCAGTAATAGTTCAAA	1617
Db	123	AATTTTCTCTTTCTATGTAAGTATTTTGCATATAAACTTGCAGTAATAGTTCAAA	182
Qy	1618	AATTAATAGTTTGTGACATTTGGCTTTCTGAGAAGAGAAATTCGAAAGTGTCAAAAATAA	1677
Db	183	AATTAATAGTTTGTGACATTTGGCTTTCTGAGAAGAGAAATTCGAAAGTGTCAAAAATAA	242
Qy	1678	AAAAAGATGAATGAAGCATATATATATTTTCAATTTTCTAGTCAACAGAGA	1737
Db	243	AAAAAGATGAATGAAGCATATATATATTTTCAATTTTCTAGTCAACAGAGA	302
Qy	1738	ATCGAAGGATCTGTTTCAAAATATTAGTAAAAATTTGAAAAATAAACTTGTCTTATTTTG	1797
Db	303	ATCGAAGGATCTGTTTCAAAATATTAGTAAAAATTTGAAAAATAAACTTGTCTTATTTTG	362
Qy	1798	TTTGCAACACACTAGTTAAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGATGT	1857
Db	363	TTTGCAACACACTAGTTAAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGATGT	422
Qy	1858	AGTTCTGTTTAAAAATTTCAAGCAAACTGGAATAATTAATCCATCTAATATGCTTTCTTT	1917
Db	423	AGTTCTGTTTAAAAATTTCAAGCAAACTGGAATAATTAATCCATCTAATATGCTTTCTTT	482
Qy	1918	CCCAAGAGTTTAAAAATTTCAAGCAAACTGGAATAATTTGGAGACAAAAGCCTTAATGA	1977
Db	483	CCCAAGAGTTTAAAAATTTCAAGCAAACTGGAATAATTTGGAGACAAAAGCCTTAATGA	542
Qy	1978	CAATGCAATTCATATATATTTTGTATAGTTTACAGTATACAGTATACAGTATACAGTAT	2037
Db	543	CAATGCAATTCATATATATTTTGTATAGTTTACAGTATACAGTATACAGTATACAGTAT	602
Qy	2038	ATGAGATGCTTGGACACAGAGTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTT	2097
Db	603	ATGAGATGCTTGGACACAGAGTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTT	662
Qy	2098	TTTCCATACATATAATGAGAGCTTCGAAAAATGGGATTCAGTCTAATCATATAAATTCAC	2157
Db	663	TTTCCATACATATAATGAGAGCTTCGAAAAATGGGATTCAGTCTAATCATATAAATTCAC	722
Qy	2158	TTATGTTTGTATATACACCTTTATCTGAATAGCTTGAAGGTAATTTTATACAAATTTTAA	2217
Db	723	TTATGTTTGTATATACACCTTTATCTGAATAGCTTGAAGGTAATTTTATACAAATTTTAA	782
Qy	2218	TAAATTTATGCTGAAACAGAGTTTGGGACCATTTGGACCATTCAGAAAGAGAGTGTAC	2277
Db	783	TAAATTTATGCTGAAACAGAGTTTGGGACCATTTGGACCATTCAGAAAGAGAGTGTAC	842
Qy	2278	TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATCTTAAAGTGTGATGATGATGATGATG	2337
Db	843	TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATCTTAAAGTGTGATGATGATGATGATG	902
Qy	2338	TGATCCGAGTACTTTGGGAAAGCCAGACAGGTGGATCTCTCTGAGCCAGAGTTTGAAGC	2397
Db	903	TGATCCGAGTACTTTGGGAAAGCCAGACAGGTGGATCTCTCTGAGCCAGAGTTTGAAGC	962
Qy	2398	CAGACTGCAACACAGTGTGAGCTGTTTCTCAATTAATTAATAAATTTAGCCAGGTGT	2457
Db	963	CAGACTGCAACACAGTGTGAGCTGTTTCTCAATTAATTAATAAATTTAGCCAGGTGT	1022
Qy	2458	GGTGTGTCACACCTGTAGTCTCCAGGCTACTCAGGAGGTGAGG-TAGTAGGATTTGTGAG	2516
Db	1023	GGTGTGTCACACCTGTAGTCTCCAGGCTACTCAGGAGGTGAGG-TAGTAGGATTTGTGAG	1082
Qy	2517	ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGGCCACCA-2560	





Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1709)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
Location/Qualifiers  
1..1709  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NOVAR2000710"  
/tissue\_type="ovary"  
/clone\_lib="NOVAR2"  
/dev\_stage="adult"  
/note="cloning vector: pME18SFL3"

FEATURES  
source  
44.5%; Score 1673.4; DB 9; Length 1709;  
Best Local Similarity 99.9%; Pred. No. 2.6e-282;  
Matches 1674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN  
Query Match  
1 GAGAAACCGAGTCACCTGTGAAAGATGGGAAATATCTCTCGGAAACTCAGTTGCCTG 60  
DB 35 GAGAAACCGAGTCACCTGTGAAAGATGGGAAATATCTCTCGGAAACTCAGTTGCCTG 94  
QY 61 GAGAGAAATCAAAGAACCCAGAAAGGAAACCCAGATGAGGAAAGAAACCGGAGGAA 120  
DB 95 GAGAGAAATCAAAGAACCCAGAAAGGAAACCCAGATGAGGAAAGAAACCGGAGGAA 154  
QY 121 ATGACTACATTTGAAAGAAACTTCAAGATCAGATGAGGAAAGGAAAGGAAAGTTTCATCC 180  
DB 155 ATGACTACATTTGAAAGAAACTTCAAGATCAGATGAGGAAAGGAAAGGAAAGTTTCATCC 214  
QY 181 ACTTCTAATCAGGAAACCGAGAAATGGCAGTGGTTCTTGAAGAGTGTGTACACTGTCAAT 240  
DB 215 ACTTCTAATCAGGAAACCGAGAAATGGCAGTGGTTCTTGAAGAGTGTGTACACTGTCAAT 274  
QY 241 AATCACAATCCCATCAGAGATCCTCTGAGCTCCAAATGATGATGGCTATGAGAACAT 300  
DB 275 AATCACAATCCCATCAGAGATCCTCTGAGCTCCAAATGATGATGGCTATGAGAACAT 334  
QY 301 GACTCCCTCACAAGGAAAGTGAGACAGTTTGAAGAAAGTTCAGAGACAGAAATATGCCCT 360  
DB 335 GACTCCCTCACAAGGAAAGTGAGACAGTTTGAAGAAAGTTCAGAGACAGAAATATGCCCT 394  
QY 361 CTTAGGACTTCTGTAGTAGGCTTGTCTCTGACCCATCAGCATGATTAAGAGTTGTG 420  
DB 395 CTTAGGACTTCTGTAGTAGGCTTGTCTCTGACCCATCAGCATGATTAAGAGTTGTG 454  
QY 421 TTTCCACACTAAATCCTCAAGCTGTTTATACCTTCCAGCAATGAGCAATGAGCAATGAGAA 480  
DB 455 TTTCCACACTAAATCCTCAAGCTGTTTATACCTTCCAGCAATGAGCAATGAGCAATGAGAA 514  
QY 481 TAGCAGACTCTGGCGAAGTTGTTTCACTCAGCAGTGCATGAAACATTTCTTCTGGCTA 540  
DB 515 TAGCAGACTCTGGCGAAGTTGTTTCACTCAGCAGTGCATGAAACATTTCTTCTGGCTA 574  
QY 541 AAGTTTAGAAATATATCTTATATATATATCTTAGGCACTCTGATATGGCAATCTCTG 600  
DB 575 AAGTTTAGAAATATATCTTATATATATATCTTAGGCACTCTGATATGGCAATCTCTG 634

QY 601 TGGCTTAGTGAATCATAGAAATTCACACAAATGACCTAAATATTTCTATGCTTTTTC 660  
DB 635 TGGCTTAGTGAATCATAGAAATTCACACAAATGACCTAAATATTTCTATGCTTTTTC 694  
QY 661 TTGTAAAGTTTGAGGACATGGAGGTGATAAAAAAATCTTTCTTAGGACAAATATGTAATA 720  
DB 695 TTGTAAAGTTTGAGGACATGGAGGTGATAAAAAAATCTTTCTTAGGACAAATATGTAATA 754  
QY 721 TGAATAAATTTCTAATCCCTCTGACTAACTGAATGAGCCCTCTTCTTAGGACAAAGAGA 780  
DB 755 TGAATAAATTTCTAATCCCTCTGACTAACTGAATGAGCCCTCTTCTTAGGACAAAGAGA 814  
QY 781 CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGAGACACACC 840  
DB 815 CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGAGACACACC 874  
QY 841 TTGTATATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGAGTCAATAGACT 900  
DB 875 TTGTATATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGAGTCAATAGACT 934  
QY 901 GATGAAATAGACTGATTTGGCAATTAAGAGTCCCAATTCACCTGACTCTGGTGTAGAT 960  
DB 935 GATGAAATAGACTGATTTGGCAATTAAGAGTCCCAATTCACCTGACTCTGGTGTAGAT 994  
QY 961 CACACACTCTGTGAGGATTTCCATCTATGAGACTTTTGTCTACATAACAGAGACTTGGTT 1020  
DB 995 CACACACTCTGTGAGGATTTCCATCTATGAGACTTTTGTCTACATAACAGAGACTTGGTT 1054  
QY 1021 TCACAACCCCTTTTATTTTAGCTAAAGCAATTTCTTTCTACTGACTTTCTTAAGTCTTTAGA 1080  
DB 1055 TCACAACCCCTTTTATTTTAGCTAAAGCAATTTCTTTCTACTGACTTTCTTAAGTCTTTAGA 1114  
QY 1081 CAAAGCTTAACCTTTTCAACCAATTCGCAATCAGACAACTTTGAAATCTACCTATGACCT 1140  
DB 1115 CAAAGCTTAACCTTTTCAACCAATTCGCAATCAGACAACTTTGAAATCTACCTATGACCT 1174  
QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTCTTAAAGCTGAAACCCATGTGCACTTTTCCA 1200  
DB 1175 GTAAGCTCTCTCTGCTTCAAGATCTTCTTAAAGCTGAAACCCATGTGCACTTTTCCA 1234  
QY 1201 TTTAATGATTTATGCTTTGCTTTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260  
DB 1235 TTTAATGATTTATGCTTTGCTTTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1294  
QY 1261 GTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320  
DB 1295 GTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1354  
QY 1321 GGTAAAGTATGCTTTGCTTTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
DB 1355 GGTAAAGTATGCTTTGCTTTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1414  
QY 1381 TTAACCAATAGCTTCCACAAATATATGTCCAAAGTCTTCAATTCCAAGCTCTGCTACCA 1440  
DB 1415 TTAACCAATAGCTTCCACAAATATATGTCCAAAGTCTTCAATTCCAAGCTCTGCTACCA 1474  
QY 1441 AATTTCAATGCAACATCTCCCATCCCAATTAACCTTTTCACTTTTGAAGTGTAACTA 1500  
DB 1475 AATTTCAATGCAACATCTCCCATCCCAATTAACCTTTTCACTTTTGAAGTGTAACTA 1534  
QY 1501 CTCATAAATCTGTGTAAGCCAGTGCAGACCTTTTGTCTTAACTGACATTTTCACTTCAAT 1560  
DB 1535 CTCATAAATCTGTGTAAGCCAGTGCAGACCTTTTGTCTTAACTGACATTTTCACTTCAAT 1594  
QY 1561 TTTTCTTTTCTATGTAATTTTGTGATATATATTTGTGATATAAATTTGAGTAATAGTCAAAAT 1620  
DB 1595 TTTTCTTTTCTATGTAATTTTGTGATATATATTTGTGATATAAATTTGAGTAATAGTCAAAAT 1654  
QY 1621 TAATAGTTTGTACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAT 1675  
DB 1655 TAATAGTTTGTACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAT 1709

FEATURES		Location/Qualifiers
source	1..1709	/organism="Homo sapiens"
		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"
ORIGIN		
Query Match		44.5%; Score 1673.4; DB 6; Length 1709;
Best Local Similarity		99.9%; Pred. No. 2.6e-282;
Matches 1674; Conservative		0; Mismatches 1; Indels 0; Gaps 0;
Qy	1	GAGAAACCGAGTCTACTGTGAAAGATGGAAATATATCTCTCGCGAAAACTCAGTTGGCTG 60
Db	35	GAGAAACCGAGTCTACTGTGAAAGATGGAAATATATCTCTCGCGAAAACTCAGTTGGCTG 94
Qy	61	GGAGAGATCAAAAGAACCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
Db	95	GGAGAGATCAAAAGAACCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGGAA 154
Qy	121	ATGACTACATTTGAAAGAAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 180
Db	155	ATGACTACATTTGAAAGAAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 214
Qy	181	ACTTCTAATCAGGAAACGAGATGGCAGTGGTCTCTGAAGAGTGTGCTACACTGTCAAT 240
Db	215	ACTTCTAATCAGGAAACGAGATGGCAGTGGTCTCTGAAGAGTGTGCTACACTGTCAAT 274
Qy	241	AATCACATCCCCATCAGAGATCTCTCCCTGAGCTCCAAATGATGAGGCTATGAGAACAT 300
Db	275	AATCACATCCCCATCAGAGATCTCTCCCTGAGCTCCAAATGATGAGGCTATGAGAACAT 334
Qy	301	GACTCCCTCAAGAGAAAGTGAAGAGTGTGAGAGAGGTGAGAGAGGTGAGAGAGGTGAGAG 360
Db	335	GACTCCCTCAAGAGAAAGTGAAGAGTGTGAGAGAGGTGAGAGAGGTGAGAGAGGTGAGAG 394
Qy	361	CTTAGGACTTCTGTTAGTAGGCTTGTCTCTGACCCATGAGCATGATTATGAAGTTGTG 420
Db	395	CTTAGGACTTCTGTTAGTAGGCTTGTCTCTGACCCATGAGCATGATTATGAAGTTGTG 454
Qy	421	TTTCCACACTAAATCTCAAGCTGCTTTATCACTTCCAGCAATGAAGCAATGAAGCA 480
Db	455	TTTCCACACTAAATCTCAAGCTGCTTTATCACTTCCAGCAATGAAGCAATGAAGCA 514
Qy	481	TAGCAGACTCGGCGAAGTGTCTCACCTGAGCAGTGCATGAACATCTCTCTGCGCTA 540
Db	515	TAGCAGACTCGGCGAAGTGTCTCACCTGAGCAGTGCATGAACATCTCTCTGCGCTA 574
Qy	541	AAGTTTGAATAATTTCTATTATATATATATATATATATATATATATATATATATATAT 600
Db	575	AAGTTTGAATAATTTCTATTATATATATATATATATATATATATATATATATATATAT 634
Qy	601	TGGCTTAGTGAAATCATAGAAATGACAAATGACCTAAATATTTCTATGTGTTTTTGC 660
Db	635	TGGCTTAGTGAAATCATAGAAATGACAAATGACCTAAATATTTCTATGTGTTTTTGC 694
Qy	661	TTGTAAGTTTGGAGCATGAGGTGATAAATAAACTTTCTTAGGCAATTAATGTAATA 720
Db	695	TTGTAAGTTTGGAGCATGAGGTGATAAATAAACTTTCTTAGGCAATTAATGTAATA 754
Qy	721	TGAAATAAATTTCTAATCCCTGACTTAATGAAATGAAATGAAATGAAATGAAATGAA 780
Db	755	TGAAATAAATTTCTAATCCCTGACTTAATGAAATGAAATGAAATGAAATGAAATGAA 814
Qy	781	CCTCAGATGAACCTGAAGACATGAATCTGCGCATGATAGGAGGAGGTGAGACACACC 840
Db	815	CCTCAGATGAACCTGAAGACATGAATCTGCGCATGATAGGAGGAGGTGAGACACACC 874
Qy	841	TTGTTATACCCCTTCCCTTTTGGAGTTTATCCCAAGTGAACGAGTGAACGAGTGAAC 900
Db	875	TTGTTATACCCCTTCCCTTTTGGAGTTTATCCCAAGTGAACGAGTGAACGAGTGAAC 934
Qy	901	GATGAAATAGACTGATTGTGGCAATAGAGTCCCAATTTCCCAATCTGACTCTGGTGTAGAT 960

Db	935	GATGAAATAGACTGATTGTGGCAATAGAGTCCCAATTTCCAACTCTGGTGTAGAT 994
Qy	961	CACACTGTCTGAGGATTCATCTATGAGACTTTTGTCTACATAACAGAGACCTTGGTT 1020
Db	995	CACACTGTCTGAGGATTCATCTATGAGACTTTTGTCTACATAACAGAGACCTTGGTT 1054
Qy	1021	TCGCAACCCCTTTATTTTATGCTTAAAGCATTTCTTTCTACTGACTTCTTAAAGTCTTTAGA 1080
Db	1055	TCGCAACCCCTTTATTTTATGCTTAAAGCATTTCTTTCTACTGACTTCTTAAAGTCTTTAGA 1114
Qy	1081	CAAAAGCTTAACTCTTTTCAACCAATTCGCAATTCAGACAACTTTGAATCTACCTATGACT 1140
Db	1115	CAAAAGCTTAACTCTTTTCAACCAATTCGCAATTCAGACAACTTTGAATCTACCTATGACT 1174
Qy	1141	GTAAGCTCTCTCTCTGCTTCAAGATCTTTCGCTCTTTTAAAGCTGAAACGATGTGCATTTCCA 1200
Db	1175	GTAAGCTCTCTCTCTGCTTCAAGATCTTTCGCTCTTTTAAAGCTGAAACGATGTGCATTTCCA 1234
Qy	1201	TTTAATGATTTATGCTCTTTCGCTTGTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
Db	1235	TTTAATGATTTATGCTCTTTCGCTTGTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1294
Qy	1261	GTGACCTGACACCTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCGAT 1320
Db	1295	GTGACCTGACACCTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCGAT 1354
Qy	1321	GGTAAGTCATGTTGGCTCAGAAATCAACCTCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
Db	1355	GGTAAGTCATGTTGGCTCAGAAATCAACCTCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1414
Qy	1381	TTACCAATAGTCTCCCAAAATATATATGTCAGAACTCTTCAATTCGAAGCTGTCTACCA 1440
Db	1415	TTACCAATAGTCTCCCAAAATATATATGTCAGAACTCTTCAATTCGAAGCTGTCTACCA 1474
Qy	1441	AAATTCAAAATGCCAATCTCTCCCATCAATTAACCTATTTTCTATTTTCTATTTTCTATTTT 1500
Db	1475	AAATTCAAAATGCCAATCTCTCCCATCAATTAACCTATTTTCTATTTTCTATTTTCTATTTT 1534
Qy	1501	CTCAATAAATCTGTGTAAGACCAAGTGCAGACCTTTTGTCTAACTGACATTTTCTTCAAT 1560
Db	1535	CTCAATAAATCTGTGTAAGACCAAGTGCAGACCTTTTGTCTAACTGACATTTTCTTCAAT 1594
Qy	1561	TTTCTTTTCTATGCTGATGATTTTCTCATATATAAATTTGAGTGAATGAGTGAATGAGTGA 1620
Db	1595	TTTCTTTTCTATGCTGATGATTTTCTCATATATAAATTTGAGTGAATGAGTGAATGAGTGA 1654
Qy	1621	TAATAGTTTGTGACTTGGCTTTTCTGAGAAGAAATTCGAAAGTGTCAAAAAT 1675
Db	1655	TAATAGTTTGTGACTTGGCTTTTCTGAGAAGAAATTCGAAAGTGTCAAAAAT 1709

RESULT 9	
AK123798	
LOCUS	AK123798 Homo sapiens cDNA FLJ41804 fis, clone NOV2000710.
DEFINITION	AK123798 Homo sapiens cDNA FLJ41804 fis, clone NOV2000710.
ACCESSION	AK123798
VERSION	AK123798.1 GI:34529425
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Moniyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Teraohima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,

QY 721 TGAATAAATTTCTAATCCCTGACTAAGTGAAGGACCTCTCTTAGGCCAAAGAGA 780  
 Db 744 TGAATAAATTTCTAATCCCTGACTAAGTGAAGGACCTCTCTTAGGCCAAAGAGA 803  
 QY 781 CTTAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGAGACACCC 840  
 Db 804 CTTAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGAGACACCC 863  
 QY 841 TTGTTATACCCCTTCCCTTTGGAGTTTATGCAAGTGCACAGGATGAGTCAATAGACT 900  
 Db 864 TTGTTATACCCCTTCCCTTTGGAGTTTATGCAAGTGCACAGGATGAGTCAATAGACT 923  
 QY 901 GATGAATAGACTGATTTGTTGGCAATAGAGTCCCAATTCGAACCTGACTCTGTTGATGAT 960  
 Db 924 GATGAATAGACTGATTTGTTGGCAATAGAGTCCCAATTCGAACCTGACTCTGTTGATGAT 983  
 QY 961 CACACACTGTCTGAGGAGTTCCATCTATGAGACTTTTGTCTACATAACAGAGACCTTGGTT 1020  
 Db 984 CACACACTGTCTGAGGAGTTCCATCTATGAGACTTTTGTCTACATAACAGAGACCTTGGTT 1043  
 QY 1021 TCCAGAACCCCTTTATTTAGCTTAAAGCAATCTTTTCTACTGACTCTTTAAAGTCTTTAGA 1080  
 Db 1044 TCCAGAACCCCTTTATTTAGCTTAAAGCAATCTTTTCTACTGACTCTTTAAAGTCTTTAGA 1103  
 QY 1081 CAAAGCTTAACTCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1140  
 Db 1104 CAAAGCTTAACTCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1163  
 QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTTGCTCTTTAAAGCTGAAACCGATGTGCACTTTCCA 1200  
 Db 1164 GTAAGCTCTCTCTGCTTCAAGATCTTTGCTCTTTAAAGCTGAAACCGATGTGCACTTTCCA 1223  
 QY 1201 TTTAATGATTTATGCTTTGCTTGAATCTCTGCTCTCTCTTAAAGTGTATTAAGTAAAG 1260  
 Db 1224 TTTAATGATTTATGCTTTGCTTGAATCTCTGCTCTCTCTTAAAGTGTATTAAGTAAAG 1283  
 QY 1261 GTGACCTGACCACTCTCAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCAGGCCAT 1320  
 Db 1284 GTGACCTGACCACTCTCAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCAGGCCAT 1343  
 QY 1321 GGTAACTCATGTTGGCTCAGAACTCAACCTCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1380  
 Db 1344 GGTAACTCATGTTGGCTCAGAACTCAACCTCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1403  
 QY 1381 TTTACCAATAGCTTCCACAAATATATGTCACAAATCTTCAATTCGAGCTCTGTCACCA 1440  
 Db 1404 TTTACCAATAGCTTCCACAAATATATGTCACAAATCTTCAATTCGAGCTCTGTCACCA 1463  
 QY 1441 AATTTCAAAATGCCAAATCTCCCAATCCCAATTTACCTATTTTCACTTTGAGGTGTAACTTA 1500  
 Db 1464 AATTTCAAAATGCCAAATCTCCCAATCCCAATTTACCTATTTTCACTTTGAGGTGTAACTTA 1523  
 QY 1501 CTTCAATAAATCTGTGAAGACAGTACAGACCTTTTGTCTTAACTGACATTTACTTCAAT 1560  
 Db 1524 CTTCAATAAATCTGTGAAGACAGTACAGACCTTTTGTCTTAACTGACATTTACTTCAAT 1583  
 QY 1561 TTTTCTTTTCTATGTTAGTATTTTCTCATATAAATTTGAGTGTAACTTGTCAAAAT 1620  
 Db 1584 TTTTCTTTTCTATGTTAGTATTTTCTCATATAAATTTGAGTGTAACTTGTCAAAAT 1643  
 QY 1621 TAATAGTTTTTGACATTTGGCTTTCTGAGAGAGAAATTCGAAAGTGTCAAAAATAAAAA 1680  
 Db 1644 TAATAGTTTTTGACATTTGGCTTTCTGAGAGAGAAATTCGAAAGTGTCAAAAATAAAAA 1703  
 QY 1681 AGATGAATGAAGCATATATATTTGTCATATTTTCAATTTTCTAGTCAACAGAGATC 1740  
 Db 1704 AGATGAATGAAGCATATATATTTGTCATATTTTCAATTTTCTAGTCAACAGAGATC 1763  
 QY 1741 GAAGGATTTCTGTTCAATATTTAGTAAATAATTTGAGAAATATGCTGCTTATATTTGTTT 1800  
 Db 1764 GAAGGATTTCTGTTCAATATTTAGTAAATAATTTGAGAAATATGCTGCTTATATTTGTTT 1823  
 QY 1801 GCAACACACTAGTTTAACTTAACTGTGACTAGTTATCTCTACCGAAGGTGATGTGATG 1860

Db 1824 GCACACACTAGTTAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1883  
 QY 1861 TTCTGGTTTTAAATTTCAAGCAAACTGGAATAATCCATCTAAATTTATGCTTTCTTTCCC 1920  
 Db 1884 TTCTGGTTTTAAATTTCAAGCAAACTGGAATAATCCATCTAAATTTATGCTTTCTTTCCC 1943  
 QY 1921 AGAAGCTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAAGCCTTAATTCGACAA 1980  
 Db 1944 AGAAGCTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAAGCCTTAATTCGACAA 2003  
 QY 1981 TGCATTCAATATATTTTTTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAGATG 2040  
 Db 2004 TGCATTCAATATATTTTTTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAGATG 2063  
 QY 2041 AGATGCTTGGACACAGAGTGTTTTGGATTTTCAGATTTATTTTGGATTTTGGAAATTTT 2100  
 Db 2064 AGATGCTTGGACACAGAGTGTTTTGGATTTTTCAGATTTATTTTGGATTTTGGAAATTTT 2123  
 QY 2101 CCATACATATAATGAGAGAGTTGGAAAATGGGATTCAGTCTAATCATAAAATTCACCTTA 2160  
 Db 2124 CCATACATATAATGAGAGAGTTGGAAAATGGGATTCAGTCTAATCATAAAATTCACCTTA 2183  
 QY 2161 TGTTCATATACACCTTATCTGAATAGCCTGAAGGTAAATTTTATACAATATTTTAAATAA 2220  
 Db 2184 TGTTCATATACACCTTATCTGAATAGCCTGAAGGTAAATTTTATACAATATTTTAAATAA 2243  
 QY 2221 TTTTATGCTGAAAACAGAGTTCGCGACATTTGGACCATCAGAAAGCAGAGTGTCACTAT 2280  
 Db 2244 TTTTATGCTGAAAACAGAGTTCGCGACATTTGGACCATCAGAAAGCAGAGTGTCACTAT 2303  
 QY 2281 TTCAAGTCACTGCTCAAAAAGTTTCAGATGTTTAAAGTGGTGTAGTCACTGAGTTCATGCCAGTGA 2340  
 Db 2304 TTCAAGTCACTGCTCAAAAAGTTTCAGATGTTTAAAGTGGTGTAGTCACTGAGTTCATGCCAGTGA 2363  
 QY 2341 TCCGAGTACTTTGGGAAGCCCAAGACAGTGTGATCTCTTGAGCCCGAGGATTTGAGGCCAG 2400  
 Db 2364 TCCGAGTACTTTGGGAAGCCCAAGACAGTGTGATCTCTTGAGCCCGAGGATTTGAGGCCAG 2423  
 QY 2401 ACTGCAACACAGTCAAGACCTGTTTCTACAATAATTAATAAAATTTAGCCAGTGTGGT 2460  
 Db 2424 ACTGCAACACAGTCAAGACCTGTTTCTACAATAATTAATAAAATTTAGCCAGTGTGGT 2483  
 QY 2461 GGTGCACACTGTGAGTCCAGGACTCTCAGGAGCTCAGGAGTGTAGGATTTGTTGAGACTG 2520  
 Db 2484 GGTGCACACTGTGAGTCCAGGACTCTCAGGAGCTCAGGAGTGTAGGATTTGTTGAGACTG 2543  
 QY 2521 GGAAGTTGAGGCTGAATCTGAGCCAGGATCTTGCCACACATTCAGCTTGGGCAACAGAG 2580  
 Db 2544 GGAAGTTGAGGCTGAATCTGAGCCAGGATCTTGCCACACATTCAGCTTGGGCAACAGAG 2603  
 QY 2581 TGAGACCTGTCTCAAAAAAATAAAAAA 2608  
 Db 2604 TGAGACCTGTCTCAAAAAAATAAAAAA 2631

RESULT 8

LOCUS CQ842314 1709 bp DNA linear PAT 02-AUG-2004  
 DEFINITION Sequence 961 from Patent EP1440981.  
 ACCESSION CQ842314  
 VERSION CQ842314.1 GI:50894101  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,  
 Yamanoto, J., Isono, Y., Nagai, K. and Irie, R.  
 TITLE Full-length human cdna  
 JOURNAL Patent: EP 1440981-A 961 28-JUL-2004;  
 Research Association for Biotechnology (JP)

Db 2478 CCACAGTCTCTTTTCCCTGAGTTTCCACAGTCTCGCAACCAATGATCGCTTCGTAT 2419  
 Qy 3473 AATTATAACTGTTCTAGATATTTGTAGCAATGTACCCCTTCCATATTTATTTTGTGTG 3532  
 Db 2418 AATTATAACTGTTCTAGATATTTGTAGCAATGTACCCCTTCCATATTTATTTTGTGTG 2359  
 Qy 3533 TAAGGCTTCTTTTAGTCAATTAATATTTTGTAGATTCATCTATGTTTAAATGTTCTATCA 3592  
 Db 2358 TAAGGCTTCTTTTAGTCAATTAATATTTTGTAGATTCATCTATGTTTAAATGTTCTATCA 2299  
 Qy 3593 GTAGTTGTACATCTTACTGTCTCAGCATATCACCATATAGATATATTAATTTGTAA 3652  
 Db 2298 GTAGTTGTACATCTTACTGTCTCAGCATATCACCATATAGATATTAATTTGTAA 2239  
 Qy 3653 TCTAATCACTGATGATATGTAGGATATTTAAGTTTGTGACATTAATGAATGAGGCTA 3712  
 Db 2238 TCTAATCACTGATGATATGTAGGATATTTAAGTTTGTGACATTAATGAATGAGGCTA 2179  
 Qy 3713 TAAATGAATAAAGCGACTACAAA 3735  
 Db 2178 TAAATGAATAAAGCGACTACAAA 2156

RESULT 7  
 BC024174  
 LOCUS Homo sapiens LOC148823, mRNA (cdna clone IMAGE:4109064).  
 DEFINITION BC024174  
 ACCESSION BC024174.1 GI:18848218  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2648)  
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schmeitz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Buetow,K.H., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 2648)  
 Strausberg,R.  
 Direct Submission  
 Submitted (19-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbe-r@mail.nih.gov](mailto:cgapbe-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>

TITLE  
 JOURNAL  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

REMARK  
 COMMENT

contact: amadan@systemsbiology.org  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.illnl.gov>  
 Series: IRL Plate: 32 Row: b Column: 10  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21687148.

FEATURES  
 Location/Qualifiers  
 source  
 1..2648  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4109064"  
 /tissue\_type="Bone marrow, acute myelogenous leukemia"  
 /clone\_lib="NIH\_MGC\_55"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"

## ORIGIN

Query Match 69.2%; Score 2604.8; DB 9; Length 2648;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2606; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 GAGAAACCGAGTCTCTGTGAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCCTG 60  
 Db 24 GAGAAACCGAGTCTCTGTGAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCCTG 83  
 Qy 61 GGAGAGAATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAGAGAGAGAAACCGCAGAA 120  
 Db 84 GGAGAGAATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAGAGAGAGAAACCGCAGAA 143  
 Qy 121 ATGACTACATTTGAAAGAAACCTTCAAGATCAAGATAAGAAAAAGCCAAAGATTTTCATCC 180  
 Db 144 ATGACTACATTTGAAAGAAACCTTCAAGATCAAGATAAGAAAAAGCCAAAGATTTTCATCC 203  
 Qy 181 ACTTTAATCAGGAAACGAGAAATGGCAGTGGTTCGAGAAAGTGTCTACATGTCATT 240  
 Db 204 ACTTTAATCAGGAAACGAGAAATGGCAGTGGTTCGAGAAAGTGTCTACATGTCATT 263  
 Qy 241 AATCACATCCCATCAGAGATCTCCCTCAGTCCCAATCATGATGCTATGAGAACATT 300  
 Db 264 AATCACATCCCATCAGAGATCTCCCTCAGTCCCAATCATGATGCTATGAGAACATT 323  
 Qy 301 GACTCCCTCAAGAGAAAGTGAGACAGTCTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTT 360  
 Db 324 GACTCCCTCAAGAGAAAGTGAGACAGTCTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTT 383  
 Qy 361 CTTAGGACTTCTGTTAGTAGGCCCTTGTCTCGACCCATGAGCATGATTAAGATTGTG 420  
 Db 384 CTTAGGACTTCTGTTAGTAGGCCCTTGTCTCGACCCATGAGCATGATTAAGATTGTG 443  
 Qy 421 TTTCCACACTAAATCTCCTCAAGTCTTTATATCACTTCCAGCAATGAAGCAATGCAGAA 480  
 Db 444 TTTCCACACTAAATCTCCTCAAGTCTTTATATCACTTCCAGCAATGAAGCAATGCAGAA 503  
 Qy 481 TAGCAGACTCTGGCGAAGTTGTTTCACTGAGCAGTGCATGAAGCAATTCCTTTCTGGCTA 540  
 Db 504 TAGCAGACTCTGGCGAAGTTGTTTCACTGAGCAGTGCATGAAGCAATTCCTTTCTGGCTA 563  
 Qy 541 AAGTTTGAATAATATATCTTTATATATATATATATATATATATATATATATATATATAT 600  
 Db 564 AAGTTTGAATAATATATCTTTATATATATATATATATATATATATATATATATATATAT 623  
 Qy 601 TGCTTTAGGTGAATCATAGAAATTCACACAATGACCTAAATATTTCTATGCTTTTTCG 660  
 Db 624 TGCTTTAGGTGAATCATATATAATGACACAATGACCTAAATATTTCTATGCTTTTTCG 683  
 Qy 661 TTGTAAGATTGTAGGACATGGAGGTGATAAAAAAACTTTTCTTAGGACAAATATATGTA 720  
 Db 684 TTGTAAGATTGTAGGACATGGAGGTGATAAAAAAACTTTTCTTAGGACAAATATATGTA 743



QY 1253 AGTAAACGGTGACCTGACCACTCAGGACACTTTCTCAGGACCTCCTCAGAGTGATATCC 1312  
 Db |||||  
 QY 4637 AGTAAACGGTGACCTGACCACTCAGGACACTTTCTCAGGACCTCCTCAGAGTGATATCC 4578  
 Db |||||  
 QY 1313 CAGGCCATGGTAAAGTCAATGCTGAGCAATCAACCTCTTTAAATATTTTACAGAAATTTG 1372  
 Db |||||  
 QY 4577 CAGGCCATGGTAAAGTCAATGCTGAGCAATCAACCTCTTTAAATATTTTACAGAAATTTG 4518  
 Db |||||  
 QY 1373 GGTGTTGGTTACCAATTAAGTCTCCACAAATATATGTCGAAGATCTTCAATTTCCAAAGCT 1432  
 Db |||||  
 QY 4517 GGTGTTGGTTACCAATTAAGTCTCCACAAATATATGTCGAAGATCTTCAATTTCCAAAGCT 4458  
 Db |||||  
 QY 1433 GCTCACCATAATTTCAATGCCAACATCTCCCAATCAATTAAGTCTTCAATTTCCAAAGCT 1492  
 Db |||||  
 QY 4457 GCTCACCATAATTTCAATGCCAACATCTCCCAATCAATTAAGTCTTCAATTTCCAAAGCT 4398  
 Db |||||  
 QY 1493 GTAATCTACTCAATTAAGTCTGTAAGACCAAGTCAAGACCACTTTCTGTAACCTGACATTT 1552  
 Db |||||  
 QY 4397 GTAATCTACTCAATTAAGTCTGTAAGACCAAGTCAAGACCACTTTCTGTAACCTGACATTT 4338  
 Db |||||  
 QY 1553 ACTTCAATTTTCTTTTCTATGTAAGTCTGTAAGACCAAGTCAAGACCACTTTCTGTAACCTGACATTT 1612  
 Db |||||  
 QY 4337 ACTTCAATTTTCTTTTCTATGTAAGTCTGTAAGACCAAGTCAAGACCACTTTCTGTAACCTGACATTT 4278  
 Db |||||  
 QY 1613 TCAAAATTAATAGTCTTGTGACANTGGCTTTCTGTAAGACCAAGTCAAGACCACTTTCTGTAACCTGACATTT 1672  
 Db |||||  
 QY 4277 TCAAAATTAATAGTCTTGTGACANTGGCTTTCTGTAAGACCAAGTCAAGACCACTTTCTGTAACCTGACATTT 4218  
 Db |||||  
 QY 1673 AATAAATAAGATGAATGAAGATATATAATGTCATTTTCTGTAACCTGACATTTTCTGTAACCTGACATTT 1732  
 Db |||||  
 QY 4217 AATAAATAAGATGAATGAAGATATATAATGTCATTTTCTGTAACCTGACATTTTCTGTAACCTGACATTT 4158  
 Db |||||  
 QY 1733 AGAATCGAAGATCTGTTCAATTAATGTAAGTCAATTAATGTAAGTCAATTAATGTAAGTCAATTAATGTAAGT 1792  
 Db |||||  
 QY 4157 AGAATCGAAGATCTGTTCAATTAATGTAAGTCAATTAATGTAAGTCAATTAATGTAAGTCAATTAATGTAAGT 4098  
 Db |||||  
 QY 1793 TTTTGTGTCACACACTAGTTAATTAACCTGTAAGTCAATTAATGTAAGTCAATTAATGTAAGTCAATTAATGTAAGT 1852  
 Db |||||  
 QY 4097 TTTTGTGTCACACACTAGTTAATTAACCTGTAAGTCAATTAATGTAAGTCAATTAATGTAAGTCAATTAATGTAAGT 4038  
 Db |||||  
 QY 1853 TGTGATGTTCTGTTGTTTAAATTTCAAGCAAACTGGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1912  
 Db |||||  
 QY 4037 TGTGATGTTCTGTTGTTTAAATTTCAAGCAAACTGGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3978  
 Db |||||  
 QY 1913 TCTTCCCAAGATGTTTAAATGATATGCAAGCTTCTTAATTTGAGAGACAAAGCTTTA 1972  
 Db |||||  
 QY 3977 TCTTCCCAAGATGTTTAAATGATATGCAAGCTTCTTAATTTGAGAGACAAAGCTTTA 3918  
 Db |||||  
 QY 1973 ATTGAATGCAATTCATATATATTTTGTATAGTTACAGTATACAGTATACAGTATACAGTATACAGTATACAGTATAC 2032  
 Db |||||  
 QY 3917 ATTGAATGCAATTCATATATATTTTGTATAGTTACAGTATACAGTATACAGTATACAGTATACAGTATACAGTATAC 3858  
 Db |||||  
 QY 2033 CTTAGATGAGATGCTTGGGACCAAGAGTGTGTTGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2092  
 Db |||||  
 QY 3857 CTTAGATGAGATGCTTGGGACCAAGAGTGTGTTGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 3798  
 Db |||||  
 QY 2093 GAATATTTCCATACATATATATGAGAGAGTGTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 2152  
 Db |||||  
 QY 3797 GAATATTTCCATACATATATGAGAGAGTGTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 3738  
 Db |||||  
 QY 2153 TTCACATTTGTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACATATTT 2212  
 Db |||||  
 QY 3737 TTCACATTTGTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACATATTT 3678  
 Db |||||  
 QY 2213 TTAATAATTTTATGCTGAAACAGAGTGTGCGACATTTGGACCATTCAGAAAGCAGAGT 2272  
 Db |||||  
 QY 3677 TTAATAATTTTATGCTGAAACAGAGTGTGCGACATTTGGACCATTCAGAAAGCAGAGT 3618  
 Db |||||  
 QY 2273 GTCACTATTTCAAGTCAAGTCTCAAAAGTTTTCAGATGTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2332  
 Db |||||  
 QY 3617 GTCACTATTTCAAGTCAAGTCTCAAAAGTTTTCAGATGTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3558  
 Db |||||

QY 2333 GCAGTGATCCGAGTACTTTGGAAAGCCAAAGCAGGTGGATCTCTTGAGCCAGGAGTTT 2392  
 Db |||||  
 QY 3557 GCAGTGATCCGAGTACTTTGGAAAGCCAAAGCAGGTGGATCTCTTGAGCCAGGAGTTT 3498  
 Db |||||  
 QY 2393 GAGGCCAGACTGCACAAACACAGTGAGACCTCGTTTCTACAAATAATTTAAAAATTTAGCCA 2452  
 Db |||||  
 QY 3497 GAGGCCAGACTGCACAAACACAGTGAGACCTCGTTTCTACAAATAATTTAAAAATTTAGCCA 3438  
 Db |||||  
 QY 2453 GGTGTTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATGTT 2512  
 Db |||||  
 QY 3437 GGTGTTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATGTT 3378  
 Db |||||  
 QY 2513 TGAGACTGGAGGTTGAGGCTCAACTGAGCCAGGATCTTGCCACCAACATTTCCAGCTTTGG 2572  
 Db |||||  
 QY 3377 TGAGACTGGAGGTTGAGGCTCAACTGAGCCAGGATCTTGCCACCAACATTTCCAGCTTTGG 3318  
 Db |||||  
 QY 2573 CAACAGAGTGAGACCTGTCTCAAAAAAAGTTTCAAGATTTTGGAGCATTTTCAAG 2632  
 Db |||||  
 QY 3317 CAACAGAGTGAGACCTGTCTCAAAAAAAGTTTCAAGATTTTGGAGCATTTTCAAG 3259  
 Db |||||  
 QY 2633 ATCTTCAGATTTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTTGACAGCATTAATC 2692  
 Db |||||  
 QY 3258 ATCTTCAGATTTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTTGACAGCATTAATC 3199  
 Db |||||  
 QY 2693 AATAGTGACTCCAGATAACTCATTTGCTGTATACACATTTTGCCTCTCTATTCAACGA 2752  
 Db |||||  
 QY 3198 AATAGTGACTCCAGATAACTCATTTGCTGTATACACATTTTGCCTCTCTATTCAACGA 3139  
 Db |||||  
 QY 2753 ATTCTTATGCCCTTTGTGTTTAAATGTCGGAAGGAAACAAATAGAAATTTTGCA 2812  
 Db |||||  
 QY 3138 ATTCTTATGCCCTTTGTGTTTAAATGTCGGAAGGAAACAAATAGAAATTTTGCA 3079  
 Db |||||  
 QY 2813 ATTCTAGAAAGTCAATCTGCAAAATATGTCAGTCTCTGTAGATATTTAGCCAAATTTTGG 2872  
 Db |||||  
 QY 3078 ATTCTAGAAAGTCAATCTGCAAAATATGTCAGTCTCTGTAGATATTTAGCCAAATTTTGG 3019  
 Db |||||  
 QY 2873 AAAAAACAAAAATTTTACTTTTCTGCTGCTTTGTAGTCTGTTTATGATATAAATPACC 2932  
 Db |||||  
 QY 3018 AAAAAACAAAAATTTTACTTTTCTGCTGCTTTGTAGTCTGTTTATGATATAAATPACC 2959  
 Db |||||  
 QY 2933 TTAATTTGTAATAAATAATTTTAAATTTGAGTAAACATCTGGAATTTATCAGAAAGGGC 2992  
 Db |||||  
 QY 2958 TTAATTTGTAATAAATAATTTTAAATTTGAGTAAACATCTGGAATTTATCAGAAAGGGC 2899  
 Db |||||  
 QY 2993 AAGCAATAGTTAATAAAGTATTTGAGTAAACATCTGGAATTTATCAGAAAGGGC 3052  
 Db |||||  
 QY 2898 AAGCAATAGTTAATAAAGTATTTGAGTAAACATCTGGAATTTATCAGAAAGGGC 2839  
 Db |||||  
 QY 3053 AATGCTCTTCTGTTGTTTCAACCAAGCCAGAGATGCTTAATCTTTCCGAGATCTAG 3112  
 Db |||||  
 QY 2838 AATGCTCTTCTGTTGTTTCAACCAAGCCAGAGATGCTTAATCTTTCCGAGATCTAG 2779  
 Db |||||  
 QY 3113 TTTTTCAGAAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTTATGAAAGCAATA 3172  
 Db |||||  
 QY 2778 TTTTTCAGAAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTTATGAAAGCAATA 2719  
 Db |||||  
 QY 3173 GAATCATCTGTTATAGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGTC 3232  
 Db |||||  
 QY 2718 GAATCATCTGTTATAGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGTC 2659  
 Db |||||  
 QY 3233 ATATAAATTTTGAACATTTTATGTTTATATTTTAACTAGTTTCTCAAGTGTGA 3292  
 Db |||||  
 QY 2658 ATATAAATTTTGAACATTTTATGTTTATATTTTAACTAGTTTCTCAAGTGTGA 2599  
 Db |||||  
 QY 3293 TTAATAAAGAAATTTGAACCTGTAAGGTTAGAGTTATAGAAATTTTGTCAATGTA 3352  
 Db |||||  
 QY 2598 TTAATAAAGAAATTTGAACCTGTAAGGTTAGAGTTATAGAAATTTTGTCAATGTA 2539  
 Db |||||  
 QY 3353 TTAACCATCTGTTAGTCACTCTTTATGAAGAGACAGAACTACATCTCTCCAGAAAGTT 3412  
 Db |||||  
 QY 2538 TTAACCATCTGTTAGTCACTCTTTATGAAGAGACAGAACTACATCTCTCCAGAAAGTT 2479  
 Db |||||  
 QY 3413 CCACAGTGCTCTTTTCCCTGAGTTTCCACAGTCTGCGCAACCAATGATCTGCTTCGTAT 3472  
 Db |||||



Sequencing vector: M13; 100%  
 Sequencing vector: plasmid; 0%  
 Chemistry: Dye-terminator ET; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 136274 bases at least Q40  
 Consensus quality: 138356 bases at least Q30  
 Consensus quality: 139182 bases at least Q20  
 Insert size: 142000; agarose-fp  
 Insert size: 140368; sum-of-coverage  
 Quality coverage: 4.64 in Q20 bases; agarose-fp  
 Quality coverage: 4.81 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 10 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 31377: contig of 31377 bp in length  
 \* 31378 31477: gap of unknown length  
 \* 31478 63011: contig of 31534 bp in length  
 \* 63012 63111: gap of unknown length  
 \* 63112 64845: contig of 1734 bp in length  
 \* 64846 64945: gap of unknown length  
 \* 64946 67606: contig of 2661 bp in length  
 \* 67607 67707: gap of unknown length  
 \* 67707 77674: contig of 9968 bp in length  
 \* 77675 77775: gap of unknown length  
 \* 77775 89257: contig of 11482 bp in length  
 \* 89257 89357: gap of unknown length  
 \* 89357 100400: contig of 11044 bp in length  
 \* 100401 100500: gap of unknown length  
 \* 100501 111915: contig of 11415 bp in length  
 \* 111916 125713: gap of unknown length  
 \* 125714 125814: contig of 13698 bp in length  
 \* 125814 141268: contig of 15455 bp in length.

FEATURES

Location/Qualifiers  
 1. 141268  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP11-115C4"  
 1. 31377  
 /note="assembly\_name:Contig10"  
 31478. 63011  
 /note="assembly\_name:Contig11"  
 63112. 64845  
 /note="assembly\_name:Contig2"  
 64946. 67606  
 /note="assembly\_name:Contig3"  
 67707. 77674  
 /note="assembly\_name:Contig4"  
 77775. 89256  
 /note="assembly\_name:Contig5"  
 89357. 100400  
 /note="assembly\_name:Contig6"  
 100501. 111915  
 /note="assembly\_name:Contig7"  
 112016. 125713  
 /note="assembly\_name:Contig8"  
 125814. 141268  
 /note="assembly\_name:Contig9"  
 clone end:SP6  
 vector\_side:left

ORIGIN

Query Match 93.3%; Score 3511.6; DB 2; Length 141268;  
 Best Local Similarity 99.3%; Pred. No. 0;

	Matches	3537;	Conservative	0;	Mismatches	25;	Indels	1;	Gaps	1;
Qy	173	TTTTCATCCACTTCTTAATCAGGAAAAACGAGAAATGCGAGTGGTCTCTGAAGAAAGTGCTTACA	232							
Db	5717	TTTCTCTTTTGTGCTTGGCAGGAAAAACGAGAAATGCGAGTGGTCTCTGAAGAAAGTGCTTACA	5658							
Qy	233	CTGTCAATTAATCAGATCCCAATCAGAGATCCCTCCCTGAGCTCCCAATGATGATGGCTATG	292							
Db	5657	CTGTCAATTAATCAGATCCCAATCAGAGATCCCTCCCTGAGCTCCCAATGATGATGGCTATG	5598							
Qy	293	AGAACATTTGACTCCCTCACAGGAAAGTGAGACAGTTTATAGAAAGTTCAGACACAAAT	352							
Db	5597	AGAACATTTGACTCCCTCACAGGAAAGTGAGACAGTTTATAGAAAGTTCAGACACAAAT	5538							
Qy	353	ATGCCCTTCTTAGGACTTCTGTCTAGTAGGCTTGTCTCTGCACCCATGAGCATGATTTATG	412							
Db	5537	ATGCCCTTCTTAGGACTTCTGTCTAGTAGGCTTGTCTCTGCACCCATGAGCATGATTTATG	5478							
Qy	413	AAGTTGTGTTTCCACACTAAATTCCTCAAGCTGCTTTATCACCTTCCAGCAATGAAGACA	472							
Db	5477	AAGTTGTGTTTCCACACTAAATTCCTCAAGCTGCTTTATCACCTTCCAGCAATGAAGACA	5418							
Qy	473	ATGCAGAAATAGCAGACTCTCGCGAAAGTTGTTCACCTGAGCAGTGCATGAACATTCCTT	532							
Db	5417	ATGCAGAAATAGCAGACTCTCGCGAAAGTTGTTCACCTGAGCAGTGCATGAACATTCCTT	5358							
Qy	533	TCTGGCTAAAGTTTAGAAATATTATTCTTATTATATATATATATATATATATATATATAT	592							
Db	5357	TCTGGCTAAAGTTTAGAAATATTATTCTTATTATATATATATATATATATATATATATAT	5298							
Qy	593	CATCTCTGTGGCTTAGTGGAATCATAGAAATGACACAAATGACCAATGACCTTAAATATCTATGT	652							
Db	5297	CATCTCTGTGGCTTAGTGGAATCATAGAAATGACACAAATGACCAATGACCTTAAATATCTATGT	5238							
Qy	653	GTTTTGTCTGTAAGTTTCAGGACATGAGAGTGTGATATAAAAAAATCTTCTTAGGACAATA	712							
Db	5237	GTTTTGTCTGTAAGTTTCAGGACATGAGAGTGTGATATAAAAAAATCTTCTTAGGACAATA	5178							
Qy	713	ATGTAAATGAAATAAATTTCTAATCCCTCCCTGACTAATCTGAATGGAACCTCTTCTTAGGC	772							
Db	5177	ATGTAAATGAAATAAATTTCTAATCCCTCCCTGACTAATCTGAATGGAACCTCTTCTTAGGC	5118							
Qy	773	CAAGAGACCTCAGATGAACCTCAAGACTGAATTCGGCCATGATAGGAAGGAGGTGA	832							
Db	5117	CAAGAGACCTCAGATGAACCTCAAGACTGAATTCGGCCATGATAGGAAGGAGGTGA	5058							
Qy	833	GACACACTTGTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTGCACAGGATGATC	892							
Db	5057	GACACACTTGTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTGCACAGGATGATC	4998							
Qy	893	ATAAGACTGATGAATAGACTGATTTGGCAATGAAGAGTCCCAATTCACACCTGACTCTG	952							
Db	4997	ATAAGACTGATGAATAGACTGATTTGGCAATGAAGAGTCCCAATTCACACCTGACTCTG	4938							
Qy	953	GTGTAGATCACACTGTCTGAGGATTCATCTATGAGACTTTGTCTAGATACACAGAGA	1012							
Db	4937	GTGTAGATCACACTGTCTGAGGATTCATCTATGAGACTTTGTCTAGATACACAGAGA	4878							
Qy	1013	CCTTGGTTTCCACAAACCCCTTTATTTTAGCTAAAGCATTTCTTCTACTGACTTCTTAAG	1072							
Db	4877	CCTTGGTTTCCACAAACCCCTTTATTTTAGCTAAAGCATTTCTTCTACTGACTTCTTAAG	4818							
Qy	1073	TCTTTAGCAAAAGCTTAACCTCTTTTCAACCAATGCGCAATCAGACAACTTTGAATCTACC	1132							
Db	4817	TCTTTAGCAAAAGCTTAACCTCTTTTCAACCAATGCGCAATCAGACAACTTTGAATCTACC	4758							
Qy	1133	TATGACTGTAACTCTCTCCCTGCTTCAAGATCTTGCCTCTTTAAAGTGAACCGATGTC	1192							
Db	4757	TATGACTGTAACTCTCTCCCTGCTTCAAGATCTTGCCTCTTTAAAGTGAACCGATGTC	4698							
Qy	1193	ACTTTCCATTAATGATTTATGCTTTGCTTTGTAATCTCTGCTCTCTCTCTCTCTCTCTCTCT	1252							
Db	4697	ACTTTCCATTAATGATTTATGCTTTGCTTTGTAATCTCTGCTCTCTCTCTCTCTCTCTCTCT	4638							

Db	133923	CTTAGATGAGATGCTTGGGACGAGAGTGTCTTTGGATTTCAGATTATTTTGGATTG	133982
QY	2093	GAATATTTCCATACATATAATGAGAGAGTTGAAATGGGATTCAGTCTAATCATATAA	2152
Db	133983	GAATATTTCCATACATATAATGAGAGAGTTGAAATGGGATTCAGTCTAATCATATAA	134042
QY	2153	TTCACTTATGTTGATATACACCTTATCTGAATAGCTCGAAGGTAAATTTTATACAAAT	2212
Db	134043	TTCACTTATGTTGATATACACCTTATCTGAATAGCTCGAAGGTAAATTTTATACAAAT	134102
QY	2213	TTAAATAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGCCATTCAGAAAGCAGA	2272
Db	134103	TTAAATAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGCCATTCAGAAAGCAGA	134162
QY	2273	GTCACTATTTCAAGTCAGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTGATCGAGTT	2332
Db	134163	GTCACTATTTCAAGTCAGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTGATCGAGTT	134222
QY	2333	GCAGTGATCCGAGTACTTTGGGAGCCCAAGACAGGTGGATCTCTTGAGCCGAGGAT	2392
Db	134223	GCAGTGATCCGAGTACTTTGGGAGCCCAAGACAGGTGGATCTCTTGAGCCGAGGAT	134282
QY	2393	GAGGCCAGACTGCACACACAGTGAGACCTCGTTTCTACAAATAATTAATAAATTAG	2452
Db	134283	GAGGCCAGACTGCACACACAGTGAGACCTCGTTTCTACAAATAATTAATAAATTAG	134342
QY	2453	GGTGTGGTGGTGACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGAT	2512
Db	134343	GGTGTGGTGGTGACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGAT	134402
QY	2513	TCGAGCTGGGAGTTGAGGCTGAACCTGAGCAGGATCTTGGCCACACATTCAGAGT	2572
Db	134403	TCGAGCTGGGAGTTGAGGCTGAACCTGAGCAGGATCTTGGCCACACATTCAGAGT	134462
QY	2573	CAACGAGTGAGACCTGTCTCAAAAATAAATAAATAAATAAATAAATAAATAAATA	2632
Db	134463	CAACGAGTGAGACCTGTCTCAAAAATAAATAAATAAATAAATAAATAAATAAATA	134521
QY	2633	ATCTTCAGATTAGGATTTTCAACCTGTACTGACCTTTAGTCAATTCAGACATTAAT	2692
Db	134522	ATCTTCAGATTAGGATTTTCAACCTGTACTGACCTTTAGTCAATTCAGACATTAAT	134581
QY	2693	AATAGTGACCTCCAGATAACTCATTTGCTGTATACACATTTTGCCTCTCTATTCA	2752
Db	134582	AATAGTGACCTCCAGATAACTCATTTGCTGTATACACATTTTGCCTCTCTATTCA	134641
QY	2753	ATCTTATGCCCTCTTGTGGTGAATTTAATGTCGGAAGGAAACAATAAGAAATTTG	2812
Db	134642	ATCTTATGCCCTCTTGTGGTGAATTTAATGTCGGAAGGAAACAATAAGAAATTTG	134701
QY	2813	ATCTAGAAAGTCATCTGTCAAAATATGTCTGCTGTAGATATTAGCCAAATTTAG	2872
Db	134702	ATCTAGAAAGTCATCTGTCAAAATATGTCTGCTGTAGATATTAGCCAAATTTAG	134761
QY	2873	AAATGACAAAAATTTTACTTTTCTGCTGCTTGTAGTGTGTTTATGATATAAATAC	2932
Db	134762	AAATGACAAAAATTTTACTTTTCTGCTGCTTGTAGTGTGTTTATGATATAAATAC	134821
QY	2933	TTATTTGTAATAAATAATTTTAAATTTGAGTAACCAATCTGGAATATTCAGAGA	2992
Db	134822	TTATTTGTAATAAATAATTTTAAATTTGAGTAACCAATCTGGAATATTCAGAGA	134881
QY	2993	AGCAATAGTTTAAATAAGATTTGATTTGGTAGAGGACGTTGAAATCCAGAGCAT	3052
Db	134882	AGCAATAGTTTAAATAAGATTTGATTTGGTAGAGGACGTTGAAATCCAGAGCAT	134941
QY	3053	AATGCTCTTGTGGTGTCCACATAAGCCACAGCAGATGTCTTAATCTTCCGAGAT	3112
Db	134942	AATGCTCTTGTGGTGTCCACATAAGCCACAGCAGATGTCTTAATCTTCCGAGAT	135001
QY	3113	TTTTTCAGCAAGCAGAGTTTAAAGAAATGTAATCTTATGTTGGTTATGAAGAA	3172
Db	135002	TTTTTCAGCAAGCAGAGTTTAAAGAAATGTAATCTTATGTTGGTTATGAAGAA	135061

RESULT 6  
AC074365/c 141268 bp DNA linear HTG 23-SEP-2000  
LOCUS Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,  
DEFINITION 10 unordered pieces.  
ACCESSION AC074365  
VERSION AC074365.5 GI:10280935  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 141268)  
Waterston,R.H.  
Direct Submission  
Submitted (29-JUL-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Sep 23, 2000 this sequence version replaced gi:9838075.  
----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
----- Project Information -----  
Center project name: H\_NH0115C04  
----- Summary Statistics -----

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----

FEATURES

Location/Qualifiers  
1..185467  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-978115"  
/clone\_lib="RPC1-11.4"

ORIGIN

Query Match 93.6%; Score 3522.2; DB 9; Length 185467;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 3544; Conservative 0; Mismatches 18; Indels 1; Gaps 1;  
QY 173 TTTCATCCACATCTTAATCAGGAAACGAGAAATGGCAGTGGTTCTGGAAGAGTGTGCTACA 232  
Db 132063 TTTCCTTTTGGCTTGGCAGGAAACGAGAAATGGCAGTGGTTCTGGAAGAGTGTGCTACA 132122  
QY 233 CTGTCATTAATACATCCCTCCATCAGAGATCTCCCTGAGCTCCAATGATGATGCTATG 292  
Db 132123 CTGTCATTAATACATCCCTCCATCAGAGATCTCCCTGAGCTCCAATGATGATGCTATG 132182  
QY 293 AGACATTTGATCTCCCTCAAGAGAAAGTGAGACAGTTTATAGAAAGGTTCAGAGACAGAAT 352  
Db 132183 AGACATTTGATCTCCCTCAAGAGAAAGTGAGACAGTTTATAGAAAGGTTCAGAGACAGAAT 132242  
QY 353 ATGCCCTTCTTAGGACTCTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATG 412  
Db 132243 ATGCCCTTCTTAGGACTCTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATG 132302  
QY 413 AAGTTGTGTTTCCACACTAAAATCCTCAAGCTGCTTTATACCTTCCAGCAATCAAGACA 472  
Db 132303 AAGTTGTGTTTCCACACTAAAATCCTCAAGCTGCTTTATACCTTCCAGCAATCAAGACA 132362  
QY 473 ATGCAGAAATAGCAGACTCTGGCGAAGTTGTTTCACTGAGCAGTGCATGAACATTCCTT 532  
Db 132363 ATGCAGAAATAGCAGACTCTGGCGAAGTTGTTTCACTGAGCAGTGCATGAACATTCCTT 132422  
QY 533 TCTGGCTAAAGTTTAGAAATATATCTATTATATATCTCTTAGCAACTCTGATATGG 592  
Db 132423 TCTGGCTAAAGTTTAGAAATATATCTATTATATATCTCTTAGCAACTCTGATATGG 132482  
QY 593 CATCTCTGTGGCTTAGGTGAATCATAGAAATGACACAAATGACCTAAATATTTCTATGT 652  
Db 132483 CATCTCTGTGGCTTAGGTGAATCATAGAAATGACACAAATGACCTAAATATTTCTATGT 132542  
QY 653 GTTTTCTGTTGTAAGTTTAGGACATGGAGTGAATAAAAAATCTTTCTTAGGACAATA 712  
Db 132543 GTTTTCTGTTGTAAGTTTAGGACATGGAGTGAATAAAAAATCTTTCTTAGGACAATA 132602  
QY 713 ATGTAAATGAAATAAATTTCTAATCCCTGACTAATCTGAATGGACCTCTCTAGGC 772  
Db 132603 ATGTAAATGAAATAAATTTCTAATCCCTGACTAATCTGAATGGACCTCTCTAGGC 132662  
QY 773 CAAAGAGACCTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGA 832  
Db 132663 CAAAGAGACCTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGA 132722  
QY 833 GACACACTGTTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGATG 892  
Db 132723 GACACACTGTTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGATG 132782  
QY 893 ATAGACTGATGAATAGACTGATTTGGCAATAAGAGTCCCAATCCCACTGACTCTG 952  
Db 132783 ATAGACTGATGAATAGACTGATTTGGCAATAAGAGTCCCAATCCCACTGACTCTG 132842  
QY 953 GTGTAGATCACACTGTCTGAGGGGATTCATCTATGAGACTTTGTCTACATAACAGAGA 1012

Db 132843 GTGTAGATCACACTGTCTGAGGATTCCTATCTATGAGACTTTGTCTACATAACAGAGA 132902  
QY 1013 CCTTGGTTTCCCAACCCCTTATTTAGCTAAAGCATTTCTTTCTACTGACTTCTTAAG 1072  
Db 132903 CCTTGGTTTCCCAACCCCTTATTTAGCTAAAGCATTTCTTTCTACTGACTTCTTAAG 132962  
QY 1073 TCTTTAGCAAAAGCTTAACTCTTTCAACCAATTCGAATCAGACAACTTTGAATCTACC 1132  
Db 132963 TCTTTAGCAAAAGCTTAACTCTTTCAACCAATTCGAATCAGACAACTTTGAATCTACC 133022  
QY 1133 TATGACCTGTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTC 1192  
Db 133023 TATGACCTGTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTC 133082  
QY 1193 ACTTTCCATTTAATGATTTATGCTTTGCTTGTAACTCTCTCTCCCTTAAATGTATATA 1252  
Db 133083 ACTTTCCATTTAATGATTTATGCTTTGCTTGTAACTCTCTCTCCCTTAAATGTATATA 133142  
QY 1253 AGTAAACGGTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCC 1312  
Db 133143 AGTAAACGGTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCC 133202  
QY 1313 CAGCCCATGTAAGTCAATGTTGGCTCAGATTCACCTCTTTAAATATTTTACAGAAATTTG 1372  
Db 133203 CAGCCCATGTAAGTCAATGTTGGCTCAGATTCACCTCTTTAAATATTTTACAGAAATTTG 133262  
QY 1373 GGTTTTGGTTTACCATAAAGTCTCCACAAATATATGTCGAAGATCTTCAATTCGAAGCT 1432  
Db 133263 GGTTTTGGTTTACCATAAAGTCTCCACAAATATATGTCGAAGATCTTCAATTCGAAGCT 133322  
QY 1433 GCTCACCAATTTCAAATGCCAATCTCCCATCCAAATTCCTATTTTCATCTTTGAGGT 1492  
Db 133323 GCTCACCAATTTCAAATGCCAATCTCCCATCCAAATTCCTATTTTCATCTTTGAGGT 133382  
QY 1493 GTATCTACTCAATAAATCTGTGAAGCAGTGAAGCAGACCTTTGCTTAACCTGACATTT 1552  
Db 133383 GTATCTACTCAATAAATCTGTGAAGCAGTGAAGCAGACCTTTGCTTAACCTGACATTT 133442  
QY 1553 ACTTCAATTTTCTTTTCTATGCTACTGGATATTTTGGCATATAAACTTCAGTGAATAGT 1612  
Db 133443 ACTTCAATTTTCTTTTCTATGTACTGGATATTTTGGCATATAAACTTCAGTGAATAGT 133502  
QY 1613 TCAAAATTAATAGTTTGGCATTTGGCTTTCTGAGAAGAAATTTGAAAGTGTCAAA 1672  
Db 133503 TCAAAATTAATAGTTTGGCATTTGGCTTTCTGAGAAGAAATTTGAAAGTGTCAAA 133562  
QY 1673 AATAAAAAAGATGAATGAAGCATATATAATTTGTCATTTTTTCAATTTCTAGTCAAC 1732  
Db 133563 AATAAAAAAGATGAATGAAGCATATATAATTTGTCATTTTTTCAATTTCTAGTCAAC 133622  
QY 1733 AGAGAATCGAAGGATTTCTGTTCAATATTTAGTAAAAATTTGAAAAATTTGCTTTATA 1792  
Db 133623 AGAGAATCGAAGGATTTCTGTTCAATATTTAGTAAAAATTTGAAAAATTTGCTTTATA 133682  
QY 1793 TTTTGTGTCACACACTAGTTAAATTAACCTGTGCTAGTTATCTCTACCGAAGTGA 1852  
Db 133683 TTTTGTGTCACACACTAGTTAAATTAACCTGTGCTAGTTATCTCTACCGAAGTGA 133742  
QY 1853 TGTGTAGTTTCTGGTTTAAATTTCAAGCAACTGGAATAATCCATCTAATTTGCTT 1912  
Db 133743 TGTGTAGTTTCTGGTTTAAATTTCAAGCAACTGGAATAATCCATCTAATTTGCTT 133802  
QY 1913 TCTTTCCCAAGAGTTTAAATTAATGATATGCTGCTTCTAAATTTGAGACAAAGCTTAA 1972  
Db 133803 TCTTTCCCAAGAGTTTAAATTAATGATATGCTGCTTCTAAATTTGAGACAAAGCTTAA 133862  
QY 1973 ATTGCAATGCAATCAATATATATTTTGTGATAGTTACAGTATACGAGTTGAGTATCC 2032  
Db 133863 ATTGCAATGCAATCAATATATATTTTGTGATAGTTACAGTATACGAGTTGAGTATCC 133922  
QY 2033 CTTTAGATGAGTCTTGGGACCAAGAGTCTTTTGGATTTTCAATTTATTTTGGATTTTG 2092

Qy	2242	TGCGC	CATTTGG	ACCACATCAG	AAAGCAG	AAGTGCTACTATTTCA	AGTCAGTGCTCAAAAAG	2301	
Db	2655	TGGCC	CATTTGG	ACCATCAG	AAAGCAG	AAGTGCTCACTATTTCA	AGTCAGTGCTCAAAAAG	2714	
Qy	2302	TTTCAG	ATGTTAG	CTGGTGAT	CGATTCATG	CCAGTGATCCG	AGTACTTTTGGGAAGCCA	2361	
Db	2715	TTTCAG	ATGTTAG	CTGGTGAT	CGATTCATG	CCAGTGATCCG	AGTACTTTTGGGAAGCCA	2774	
Qy	2362	AGACAG	TGGATCTCT	TTAGCC	CAGGAGTTT	GAGGCCAGACTG	CCAAACACAGTGGAGCC	2421	
Db	2775	AGACAG	TGGATCTCT	TTAGCC	CAGGAGTTT	GAGGCCAGACTG	CCAAACACAGTGGAGCC	2834	
Qy	2422	TGCTTT	CTCAAA	TAAATA	TTAAAA	ATTAGCAG	TGTGTGGTGCGCACTCTGTAGTCCCG	2481	
Db	2835	TGCTTT	CTCAAA	TAAATA	TTAAAA	ATTAGCAG	TGTGTGGTGCGCACTCTGTAGTCCCG	2894	
Qy	2482	GTACT	CAGGAG	CTGAG	TAGTAGGAT	TGTTT	GAGACTGGGAGTTTGAGGCTGAACTGAG	2541	
Db	2895	GTACT	CAGGAG	CTGAG	TAGTAGGAT	TGTTT	GAGACTGGGAGTTTGAGGCTGAACTGAG	2955	
Qy	2542	CCAGG	ATCTTGC	CAACCA	CATTTCC	AGCTTGGG	CAACAGAGTGAGACCTCTCTC-----AA	2596	
Db	2955	CCAGG	ATCTTGC	CAACCA	CATTTCC	AGCTTGGG	CAACAGAGTGAGACCTCTCTCAAAAAA	3014	
Qy	2597	AAAAA	AAAAA	AGTTT	CAGATTTT	TGAGCA	ATTTTCAGATCTTTCAGATAGGATTTTCAAC	2656	
Db	3015	AAAAA	AAAAA	AGTTT	CAGATTTT	TGAGCA	ATTTTCAGATCTTTCAGATAGGATTTTCAAC	3074	
Qy	2657	CTGTACT	GACCTTT	TAGTCAT	TGCACAG	CAATTAATCA	ATAGTGGGACTCCAGATACTCA	2716	
Db	3075	CTGTACT	GACCTTT	TAGTCAT	TGCACAG	CAATTAATCA	ATAGTGGGACTCCAGATACTCA	3134	
Qy	2717	TTTCG	TGTATAC	ATTTTGC	CTCTATT	CAAGAA	TTCTTATGCGCTCTTGTGGTGAT	2776	
Db	3135	TTTCG	TGTATAC	ATTTTGC	CTCTATT	CAAGAA	TTCTTATGCGCTCTTGTGGTGAT	3194	
Qy	2777	TTTAA	TGTGCG	GAAGG	AAACAT	PAGAAA	TTTTTGCAATTCAGAAAAGTCATTCTGTCAA	2836	
Db	3195	TTTAA	TGTGCG	GAAGG	AAACAT	PAGAAA	TTTTTGCAATTCAGAAAAGTCATTCTGTCAA	3254	
Qy	2837	AAATG	TCAG	TCCTGT	AGATAT	TAGCCAA	TTTTTAGGAAATGCAAAA	TTTTTACTTTT	2896
Db	3255	AAATG	TCAG	TCCTGT	AGATAT	TAGCCAA	TTTTTAGGAAATGCAAAA	TTTTTACTTTT	3314
Qy	2897	CGTCTG	CCCTTGT	AGCTGTTT	TATGAT	TAAATAC	CTTATTTGTAATATAAATTTTAA	2956	
Db	3315	CGTCTG	CCCTTGT	AGCTGTTT	TATGAT	TAAATAC	CTTATTTGTAATATAAATTTTAA	3374	
Qy	2957	ATTTG	AGTAA	CAATCTG	GAATAT	CAGAG	GGGCAAGCAATAGGTTAATAACAGTAT	3016	
Db	3375	ATTTG	AGTAA	CAATCTG	GAATAT	CAGAG	GGGCAAGCAATAGGTTAATAACAGTAT	3434	
Qy	3017	TGAT	TGGTAG	AGGAAG	CTTGAA	ATCCAG	AGCATCAATGTCCTCTGGTGGTTCA	3076	
Db	3435	TGAT	TGGTAG	AGGAAG	CTTGAA	ATCCAG	AGCATCAATGTCCTCTGGTGGTTCA	3494	
Qy	3077	AGCCA	CACAG	ATGCTT	TAATCTT	CCGAG	ATCTAGTTTTTTCAGAAAGCAGGATTTAAG	3136	
Db	3495	AGCCA	CACAG	ATGCTT	TAATCTT	CCGAG	ATCTAGTTTTTTCAGAAAGCAGGATTTAAG	3554	
Qy	3137	AAATG	TAACT	ATCTTAT	TGTTAG	GAACAA	TAGAACTATTCGCTGTATAAGTGCTTT	3196	
Db	3555	AAATG	TAACT	ATCTTAT	TGTTAG	GAACAA	TAGAACTATTCGCTGTATAAGTGCTTT	3614	
Qy	3197	TTAA	CGTG	TAAATTTT	TGGAAG	CTTACTTTT	TATGCAATATAAATTTTGAACATTTTACA	3256	
Db	3615	TTAA	CGTG	TAAATTTT	TGGAAG	CTTACTTTT	TATGCAATATAAATTTTGAACATTTTACA	3674	
Qy	3257	TTGTT	TATATTTT	TAACT	CAGTTT	TACTCA	AGTGATTTATACAAGAAAATGTAAACAC	3316	
Db	3675	TTGTT	TATATTTT	TAACT	CAGTTT	TACTCA	AGTGATTTATACAAGAAAATGTAAACAC	3734	
Qy	3317	TGTA	GGGTAG	AGTTAT	TAGAA	NTTTTGT	CAAAATGTTATCCACCATGTAGTCACTCTTTA	3376	

Db	3735	TGTAAGGGTAGAGTTATAAGAAATTTGTCAAATGTATTACCCATGTAGTCACCTCCTTA	3794
Qy	3377	TGAAGACAGAAACACGTCATCCTCCGAGAAAGTTCCACAGTGCTCCCTTTTCCCTGAGT	3436
Db	3795	TGAAGACAGAAACACGTCATCCTCCGAGAAAGTTCCACAGTGCTCCCTTTTCCCTGAGT	3854
Qy	3437	TTCCACGAGTCTCGGCAACCAATCATCTGCTTCGTATAATTAACCTGTTCTAGATATTG	3496
Db	3855	TTCCACGAGTCTCGGCAACCAATCATCTGCTTCGTATAATTAACCTGTTCTAGATATTG	3914
Qy	3497	TAGCAATGTACCCCTTTCCATATTTATTTTGTGTGTGAAGGCTCTTTTAGTCATTATAA	3556
Db	3915	TAGCAATGTACCCCTTTCCATATTTATTTTGTGTGTGAAGGCTCTTTTAGTCATTATAA	3974
Qy	3557	TATTTTGGAGATTCATCTATGTTTAAATGTTCTATCAGTAGTGTCATCTTACTCTGCTC	3616
Db	3975	TATTTTGGAGATTCATCTATGTTTAAATGTTCTATCAGTAGTGTCATCTTACTCTGCTC	4034
Qy	3617	AGCATATCACCATATAGATATACTATAAATTTGTTAATCTAAATCACTGATGGATATGAGG	3676
Db	4035	AGCATATCACCATATAGATATACTATAAATTTGTTAATCTAAATCACTGATGGATATGAGG	4094
Qy	3677	ATATTTAAGTTTTTGACATTATGAATAAAGTGCGCTATAAATG	3718
Db	4095	ATATTTAAGTTTTTGACATTATGAATAAAGTGCGCTATAAATG	4136

RESULT 5  
AL606804

AL606804	AL606804	185467 bp	DNA	linear	PRI 23-OCT-2002
LOCUS	Human DNA sequence from clone RP11-978115 on chromosome 1, complete sequence.				
DEFINITION	Human DNA sequence from clone RP11-978115 on chromosome 1, complete sequence.				
ACCESSION	AL606804				
VERSION	AL606804.11	GI:24366459			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 185467)				
AUTHORS	Almeida, J.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquest@sanger.ac.uk				
COMMENT	On Oct 24 2002 this sequence version replaced gi:18121511.				

## COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/c\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/c_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1> RP11-978I15 is from the library RPCI-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

Matches 3617; Conservative 0; Mismatches 20; Indels 5; Gaps 1;			
Qy	82	AAGAAAGAAACCCAGATGAGAAAGAAACCGGAGGAAATGACTACATTTGAAAGAAA	141
Dd	495	AAGATGGAAATTTATCTCTCGAAAACTCAGGAGGAAATGACTACATTTGAAAGAAA	554
Qy	142	CTTCAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTATCATCAGGAAACGAG	201
Dd	555	CTTCAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTATCATCAGGAAACGAG	614
Qy	202	AATGCACTGGTTCTGAGAAAGTGTGTACACTGTCTATTAATCACATCCGCCATCAGAGA	261
Dd	615	AATGCACTGGTTCTGAGAAAGTGTGTACACTGTCTATTAATCACATCCGCCATCAGAAA	674
Qy	262	TCCTCCCTGAGCTCCATGATGATGGCTATGAGAACTATGTACTCCCTCACAAAGAAAGTG	321
Dd	675	TCCTCCCTGAGCTCCATGATGATGGCTATGAGAACTATGTACTCCCTCACAAAGAAAGTG	734
Qy	322	AGACAGTTTATAGAAAGGTTCAGACAGAAATATGCCCTTCTTAGGACTCTGTGTAGTAGG	381
Dd	735	AGACAGTTTATAGAAAGGTTCAGACAGAAATATGCCCTTCTTAGGACTCTGTGTAGTAGG	794
Qy	382	CCTGTGCTCCGACCCATGAGCATGATTAATGAAGTTGTGTTTCCACACTAAAATCCTCAA	441
Dd	795	CCTGTGCTCCGACCCATGAGCATGATTAATGAAGTTGTGTTTCCACACTAAAATCCTCAA	854
Qy	442	GCTGCTTTATCACCTTCCAGCAATGAAGCAATGACAGATAGAGACTCTGGCGAAGTTG	501
Dd	855	GCTGCTTTATCACCTTCCAGCAATGAAGCAATGACAGATAGAGACTCTGGCGAAGTTG	914
Qy	502	TTCACTGAGCAGTGCATGAAACATTCCTTCTGCGCTTAAAGTTTGAAGATATATCTTA	561
Dd	915	TTCACTGAGCAGTGCATGAAACATTCCTTCTGCGCTTAAAGTTTGAAGATATATCTTA	974
Qy	562	TTATATATCCTTAGGCAACTCTGTATATGTGSCATCTCTGTGGCTTAGGTGAATCATAGA	621
Dd	975	TTATATATCCTTAGGCAACTCTGTATATGTGSCATCTCTGTGGCTTAGGTGAATCATAGA	1034
Qy	622	AATTGACACATGACCTTAAATATTTCTATGTGTTTCTGTTGTTTGAAGTTTGAAGCAATGG	681
Dd	1035	AATTGACACATGACCTTAAATATTTCTATGTGTTTCTGTTGTTTGAAGTTTGAAGCAATGG	1094
Qy	682	AGGTGATAAATAAATCTTTCTAGCACAAATGTAATAAATGAAATAAATTTCTAATCCC	741
Dd	1095	AGGTGATAAATAAATCTTTCTAGCACAAATGTAATAAATGAAATAAATTTCTAATCCC	1154
Qy	742	CCTGACTAACTGAATGACCCCTCTTCTAGCCAAAGAGACCTCAGATGAACCTGAAAGAC	801
Dd	1155	CCTGACTAACTGAATGACCCCTCTTCTAGCCAAAGAGACCTCAGATGAACCTGAAAGAC	1214
Qy	802	TGAATTCGGCCATGATAGGAGGAGGTGAGACACACCTTGTATACCCCTTCCCTTTT	861
Dd	1215	TGAATTCGGCCATGATAGGAGGAGGTGAGACACACCTTGTATACCCCTTCCCTTTT	1274
Qy	862	GGAGTTTATGCAAGTGAACAGGATGAGTCATAAGACTGATGAATAGACTGATTTGG	921
Dd	1275	GGAGTTTATGCAAGTGAACAGGATGAGTCATAAGACTGATGAATAGACTGATTTGG	1334
Qy	922	CAATAGAGTCCAAATCCAACTGACTCTGGTGTAGATCACACACTGTCTGAGGATTC	981
Dd	1335	CAATAGAGTCCAAATCCAACTGACTCTGGTGTAGATCACACACTGTCTGAGGATTC	1394
Qy	982	CATCTATGAGACTTTGTACATAACAGAGACCTTGGTTTCCACACCCCTTTATTTAG	1041
Dd	1395	CATCTATGAGACTTTGTACATAACAGAGACCTTGGTTTCCACACCCCTTTATTTAG	1454
Qy	1042	CTAAGCAATCTTTTCTACTGACTTCTTAAAGTCTTTAGACAAAGCTTAACTCTTCAACC	1101
Dd	1455	CTAAGCAATCTTTTCTACTGACTTCTTAAAGTCTTTAGACAAAGCTTAACTCTTCAACC	1514
Qy	1102	AATTGCCAATCAGACAACTTTGAATCTACCTATGACCTGTAGCTCTCTCTGCTTCAA	1161
Dd	1515	AATTGCCAATCAGACAACTTTGAATCTACCTATGACCTGTAGCTCTCTCTGCTTCAA	1574

Qy	1162	GATCTTCCCTCTTTAAGCTGAACCGATGTGCACCTTTCCATTTAATGATTTATGCTTTGC	1221
Dd	1575	GATCTTCCCTCTTTAAGCTGAACCGATGTGCACCTTTCCATTTAATGATTTATGCTTTGC	1634
Qy	1222	TTGTAACTCTCTCTCCCTAAAATGTATAAAGTAAACGGTGACCTGACCACTCAGGCA	1281
Dd	1635	TTGTAACTCTCTCTCCCTAAAATGTATAAAGTAAACGGTGACCTGACCACTCAGGCA	1694
Qy	1282	CACCTTCTCAGGACCTCCTCAGAGTGTATCCAGGCCATGTAAGTCACTGTTGCTCAGA	1341
Dd	1695	CACCTTCTCAGGACCTCCTCAGAGTGTATCCAGGCCATGTAAGTCACTGTTGCTCAGA	1754
Qy	1342	ATCAACCTCTTTAAATAATTTTACAGAAATTTGGGTTTTGGTTTACCAATAAGTCTCCAAA	1401
Dd	1755	ATCAACCTCTTTAAATAATTTTACAGAAATTTGGGTTTTGGTTTACCAATAAGTCTCCAAA	1814
Qy	1402	TATATGTCGAAGATCTTCAATTCGAAGCTGTCTACCAAAATTTCAATGCGCAACATCTC	1461
Dd	1815	TATATGTCGAAGATCTTCAATTCGAAGCTGTCTACCAAAATTTCAATGCGCAACATCTC	1874
Qy	1462	CCCATCCAATTTACTATTTCTATCTTTGAGGTGTAATCTCAATAAATCTGTGTAAGACC	1521
Dd	1875	CCCATCCAATTTACTATTTCTATCTTTGAGGTGTAATCTCAATAAATCTGTGTAAGACC	1934
Qy	1522	AGTGACGAGACCTTTGCTAACTGACATTTACTTCAATTTTCTTTTCTATGTACTGG	1581
Dd	1935	AGTGACGAGACCTTTGCTAACTGACATTTACTTCAATTTTCTTTTCTATGTACTGG	1994
Qy	1582	ATATTTTGCATATAACTTGCAGTAAATAGTTCAAAAATTTAATAGTTTGTGACATTTGGCT	1641
Dd	1995	ATATTTTGCATATAACTTGCAGTAAATAGTTCAAAAATTTAATAGTTTGTGACATTTGGCT	2054
Qy	1642	TTTCTCAGAGAGAGAAATGAAAGTGTACAAAAATGAAAGATGAAATGAAGCATATAT	1701
Dd	2055	TTTCTCAGAGAGAGAAATGAAAGTGTACAAAAATGAAAGATGAAATGAAGCATATAT	2114
Qy	1702	AATTGCAATTTTCTCAATTTCTAGTCAACAGAGAAATCGAAGGATCTGTTCAAAATATT	1761
Dd	2115	AATTGCAATTTTCTCAATTTCTAGTCAACAGAGAAATCGAAGGATCTGTTCAAAATATT	2174
Qy	1762	AGTAAAAATGAAATTAATTAATTTGCTTATATTTTGTGTCACACACACTAGTTAATTA	1821
Dd	2175	AGTAAAAATGAAATTAATTAATTTGCTTATATTTTGTGTCACACACACTAGTTAATTA	2234
Qy	1822	CCTGTGACTAGTTTATCTTACCGAAGTGTGATCTGTAGTTTCTGTTTAAATTTCAAGC	1881
Dd	2235	CCTGTGACTAGTTTATCTTACCGAAGTGTGATCTGTAGTTTCTGTTTAAATTTCAAGC	2294
Qy	1882	AAACTGGAAAAATTAATCCATCTAATTAATGCTTTCTTTCCCAAGAGTTTAAATGATG	1941
Dd	2295	AAACTGGAAAAATTAATCCATCTAATTAATGCTTTCTTTCCCAAGAGTTTAAATGATG	2354
Qy	1942	CCAGCTTCCCTAATTTGGAGACAAAGCCCTTAATTTGACATGCAATTCATTTATTTT	2001
Dd	2355	CCAGCTTCCCTAATTTGGAGACAAAGCCCTTAATTTGACATGCAATTCATTTATTTT	2414
Qy	2002	TGTATAGTTTACAGTATACAGATTGAGTATCCCTTTAGATGAGATGCTTGGGACAGAGTG	2061
Dd	2415	TGTATAGTTTACAGTATACAGATTGAGTATCCCTTTAGATGAGATGCTTGGGACAGAGTG	2474
Qy	2062	TTTTGGAATTCAGATTTATTTTGGATTTTGGATTTTGGATTTTCCATATATAATGAGAGT	2121
Dd	2475	TTTTGGAATTCAGATTTATTTTGGATTTTGGATTTTGGATTTTCCATATATAATGAGAGT	2534
Qy	2122	TGAAAAATCGGATTCAGGCTAATCATAAAAATCTACTTATGTTGATATACACCTTATCT	2181
Dd	2535	TGAAAAATCGGATTCAGGCTAATCATAAAAATCTACTTATGTTGATATACACCTTATCT	2594
Qy	2182	GAATAGCCTGGAAGGTAATTTTATACAAATTTTAAATAATTTTATGCCCTGAAACAGAGTT	2241
Dd	2595	GAATAGCCTGGAAGGTAATTTTATACAAATTTTAAATAATTTTATGCCCTGAAACAGAGTT	2654



Db 2835 TCGTTTCTACAAATTAATTAATAATTTAGCCAGGTGGTGGTGCACACTGTAGTCCAG 2894  
 Qy 2482 GTACTCAGAGGCTGAGGTAGTGGATTTGAGACTGGGAGTTGAGGCTGGAAGTCTG 2541  
 Db 2895 GTACTCAGAGGCTGAGGTAGTGGATTTGAGACTGGGAGTTGAGGCTGGAAGTCTG 2954  
 Qy 2542 CCAGATCTTGGCCACCACTTCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTC-----AA 2596  
 Db 2955 CCAGATCTTGGCCACCACTTCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTC-----AA 3014  
 Qy 2597 AAAAAAAGTTTTCAGATTTTGGAGCACTTTCAGATCTTTCAGATTTAGGGATTTTCAAC 2656  
 Db 3015 AAAAAAAGTTTTCAGATTTTGGAGCACTTTCAGATCTTTCAGATTTAGGGATTTTCAAC 3074  
 Qy 2657 CTGTACTGACCTTTTGTAGTCAATGCAAGCAATTAATCAATGAGTGGACTCCAGATCACTCA 2716  
 Db 3075 CTGTACTGACCTTTTGTAGTCAATGCAAGCAATTAATCAATGAGTGGACTCCAGATCACTCA 3134  
 Qy 2717 TTTGCTGTATACAGATTTTGGCTCTCTATTCAGCAATTTCTATGCTTATGCTTGTGGTAT 2776  
 Db 3135 TTTGCTGTATACAGATTTTGGCTCTCTATTCAGCAATTTCTATGCTTATGCTTGTGGTAT 3194  
 Qy 2777 TTTAATGTGCGAAGGAAACAAATAGAAATTTTGAATTTCTAGAAAAGTCAATTCGTCAA 2836  
 Db 3195 TTTAATGTGCGAAGGAAACAAATAGAAATTTTGAATTTCTAGAAAAGTCAATTCGTCAA 3254  
 Qy 2837 AATATGTGCTGCTGTAGTATTTAGCCAAATTTAGGAAATGACAAAATTTTACTTTT 2896  
 Db 3255 AATATGTGCTGCTGTAGTATTTAGCCAAATTTAGGAAATGACAAAATTTTACTTTT 3314  
 Qy 2897 CGTCTGCTTGTAGTCTTTTATGATTAATACCTTTATTTGTAATAAAATTAATTTA 2956  
 Db 3315 CGTCTGCTTGTAGTCTTTTATGATTAATACCTTTATTTGTAATAAAATTAATTTA 3374  
 Qy 2957 ATTTGAGTAACTCTGGAATTTATCAGAAAGGCGCAAGCAATAGTGTAAATAAACAGAT 3016  
 Db 3375 ATTTGAGTAACTCTGGAATTTATCAGAAAGGCGCAAGCAATAGTGTAAATAAACAGAT 3434  
 Qy 3017 TGATTTGTTAGAGGAAAGTGTAAATCCAAAGAGCATCAATGTCTTCGTGGTTCACCA 3076  
 Db 3435 TGATTTGTTAGAGGAAAGTGTAAATCCAAAGAGCATCAATGTCTTCGTGGTTCACCA 3494  
 Qy 3077 ACCACAGAGATGCTTAACTTTTCCGAGATCTAGTCTTTCAGCAAGACGAGATTTAG 3136  
 Db 3495 AGCCACAGAGATGCTTAACTTTTCCGAGATCTAGTCTTTCAGCAAGACGAGATTTAG 3554  
 Qy 3137 AAATGTAACTATCTTATGTTGTTATGAAGAAACAAATAGAAATCAATGCTGTATAGTCTTT 3196  
 Db 3555 AAATGTAACTATCTTATGTTGTTATGAAGAAACAAATAGAAATCAATGCTGTATAGTCTTT 3614  
 Qy 3197 TTAACCTGTAATTTTGTGAAGCTTATCTTTTATGCAATTAATAATTTGAACTTTTACA 3256  
 Db 3615 TTAACCTGTAATTTTGTGAAGCTTATCTTTTATGCAATTAATAATTTGAACTTTTACA 3674  
 Qy 3257 TTGTTTATTTTATGCTTTTACTCAAGTGTGATTTATATACAGAAATGTAACCCAC 3316  
 Db 3675 TTGTTTATTTTATGCTTTTACTCAAGTGTGATTTATATACAGAAATGTAACCCAC 3734  
 Qy 3317 TGTAGGGGTAGAGTTTATAGAAATTTTGTCAAATGTTATTCACCCATGTAGTCACTCTTA 3376  
 Db 3735 TGTAGGGGTAGAGTTTATAGAAATTTTGTCAAATGTTATTCACCCATGTAGTCACTCTTA 3794  
 Qy 3377 TGAAGAGACAGACGATCATCTCCAGAAAGTTCCACAGTGTCTTTTCCCTGAGT 3436  
 Db 3795 TGAAGAGACAGACGATCATCTCCAGAAAGTTCCACAGTGTCTTTTCCCTGAGT 3854  
 Qy 3437 TTCACAGTCTTGGCAACCAATCATCTGCTTGTATTAATTTATTAATGTTCTAGATTTTG 3496  
 Db 3855 TTCACAGTCTTGGCAACCAATCATCTGCTTGTATTAATTTATTAATGTTCTAGATTTTG 3914  
 Qy 3497 TAGCAATGTACCCCTTTCCATATTTATTTTGTGTGTGTAAGGCTTCTTTTGTAGTCAATTA 3556

Db 3915 TAGCAATGTACCCCTTTCCATATTTTATTTTGTGTGTAAGGCTTCTTTTGTAGTCAATTA 3974  
 Qy 3557 TATTTTTCAGATTCATCTATCTTTAATCTTATCAGTAGTGTGATCACTTACTTGTCTC 3616  
 Db 3975 TATTTTTCAGATTCATCTATCTTTAATCTTATCAGTAGTGTGATCACTTACTTGTCTC 4034  
 Qy 3617 AGCATATCACCATATAGATATATCTAATTTTCTTAATCTAATCACTGTGATATGTAGG 3676  
 Db 4035 AGCATATCACCATATAGATATATCTAATTTTCTTAATCTAATCACTGTGATATGTAGG 4094  
 Qy 3677 ATATTAAAGTTTTCAGATTTGATTAATAAGTGGCTATATAATG 3718  
 Db 4095 ATATTAAAGTTTTCAGATTTGATTAATAAGTGGCTATATAATG 4136

RESULT 4  
 AK124520 4136 bp mRNA linear PRI 09-SEP-2003  
 LOCUS Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.  
 DEFINITION AK124520  
 ACCESSION AK124520  
 VERSION AK124520.1 GI:34530324  
 KEYWORDS Oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,  
 Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,  
 Kamihara, K., Katsuma, N., Sato, K., Tanikawa, M., Yamazaki, M.,  
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,  
 Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,  
 Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,  
 Kikuchi, H., Murakawa, K., Kaneshiro, K., Takahashi-Fujii, A.,  
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
 Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 4136)  
 Isogai, T. and Yamamoto, J.  
 Direct Submission  
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB,  
 HRI, and Biotechnology Center, National Institute of Technology and  
 Evaluation; clone selection for full insert sequencing: HRI and  
 RAB; annotation: HRI and RAB.

FEATURES  
 Location/Qualifiers  
 1..4136  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="BRACE3002390"  
 /tissue\_type="cerebellum"  
 /clone\_lib="BRACE3"  
 /note="cloning vector: pMB18SFL3"  
 136..537  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAC85873.1"  
 /db\_xref="GI:34530325"  
 /translation="MVMLEDSSNSTCGARNVCELVFLSVLGCQSKERKQMTQQAG  
 RMLRAGREASSETNPEGLTLDLPQSPFLPAARGWQEHFVRCNVLPHPAFSWCFQG  
 NFPKWSGTQLNRNVTKRWELIICSENSGRK"  
 CDS  
 ORIGIN

Query Match 95.4%; Score 3590; DB 9; Length 4136;  
 Best Local Similarity 99.3%; Pred. No. 0;



QY	262	TCTCCCTGAGCTCCAATGATCATGGCTATAGAACATTGACTCCCTCACAGGNAAGTG	321
DB	675	TCCTCCCTGAGCTCCAATGATCATGGCTATAGAACATTGACTCCCTCACAGGAAAGTG	734
QY	322	AGACAGTTTATAGAGAAAGGTCAAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGTAGG	381
DB	735	AGACAGTTTATAGAGAAAGGTCAAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGTAGG	794
QY	382	CCTTGTTCCTGCACCATTGAGATGAATATGAAGTTGTGTTTCCACACATAAATCCTCAA	441
DB	795	CCTTGTTCCTGCACCATTGAGATGAATATGAAGTTGTGTTTCCACACATAAATCCTCAA	854
QY	442	GCTGCTTTATCACCTTCAGACAAATGAAGACAATGAGAATAGCAGACTCTGGCGAAGTTG	501
DB	855	GCTGCTTTATCACCTTCAGACAAATGAAGACAATGAGAATAGCAGACTCTGGCGAAGTTG	914
QY	502	TTACCCCTGAGCAGTGCAATGAACATTCCTTCTGGCTAAAAGTTTAGAAAAATATTATCTTA	561
DB	915	TTACCCCTGAGCAGTGCAATGAACATTCCTTCTGGCTAAAAGTTTAGAAAAATATTATCTTA	974
QY	562	TTATATATCCCTTAGGCAACTCTGATATGTGGCATCTCTGTGGCTTAGGTGAAATCATAGA	621
DB	975	TTATATATCCCTTAGGCAACTCTGATATGTGGCATCTCTGTGGCTTAGGTGAAATCATAGA	1034
QY	622	AATTGACAAATGACTAAAATATTCTATATGTTTTGTTGTTGTAAGTTTTCAGACACATGG	681
DB	1035	AATTGACAAATGACTAAAATATTCTATATGTTTTGTTGTTGTAAGTTTTCAGACACATGG	1094
QY	682	AGGTGATAAAAAAACTTTCTTAGGCAATAATGTAAAATGAAAAATTAATTTCTTAATCCC	741
DB	1095	AGGTGATAAAAAAACTTTCTTAGGCAATAATGTAAAATGAAAAATTAATTTCTTAATCCC	1154
QY	742	CCTGACTAACTGAATGAGACCTCTTCTAGGCCAAAGAGACCTCAGATGAACCTGAAAGAC	801
DB	1155	CCTGACTAACTGAATGAGACCTCTTCTAGGCCAAAGAGACCTCAGATGAACCTGAAAGAC	1214
QY	802	TGAATTTCTGGCCATTGATAGGAAGGAGGTGAGACACCTTGTATTACCCCTTCCCTTTT	861
DB	1215	TGAATTTCTGGCCATTGATAGGAAGGAGGTGAGACACCTTGTATTACCCCTTCCCTTTT	1274
QY	862	GGAGTTTATGCAAAAGTGACAGGATGAGTCATAAGACTGATGAAATAGACTGATTTGTGG	921
DB	1275	GGAGTTTATGCAAAAGTGACAGGATGAGTCATAAGACTGATGAAATAGACTGATTTGTGG	1334
QY	922	CAATAAGAGTCCCAATTTCCAACTGACTCTGGTGTAGATCACACACTGTCTGAGGGATTC	981
DB	1335	CAATAAGAGTCCCAATTTCCAACTGACTCTGGTGTAGATCACACACTGTCTGAGGGATTC	1394
QY	982	CATCTATGAGACTTTGTCTACATAACAGAGACCTTGTGTTTCCACAAACCCCTTTATTTTAG	1041
DB	1395	CATCTATGAGACTTTGTCTACATAACAGAGACCTTGTGTTTCCACAAACCCCTTTATTTTAG	1454
QY	1042	CTAAGCAATTTCTTTCTACTGACTTCTTAAGTCTTTTAGACAAAGCTTTAACTCTTTCAACC	1101
DB	1455	CTAAGCAATTTCTTTCTACTGACTTCTTAAGTCTTTTAGACAAAGCTTTAACTCTTTCAACC	1514
QY	1102	AATTGCCAATCAGACAAACTTTGAAATCTACCTATGACTGTAGCTCTCTCTGCTTCAA	1161
DB	1515	AATTGCCAATCAGACAAACTTTGAAATCTACCTATGACTGTAGCTCTCTCTGCTTCAA	1574
QY	1162	GATCTTCCCTCTTTAAGCTGAAACCGATGTGCACCTTCCATTTAATGATTTATCTCTTTGC	1221
DB	1575	GATCTTCCCTCTTTAAGCTGAAACCGATGTGCACCTTCCATTTAATGATTTATCTCTTTGC	1634
QY	1222	TTGTAACTCCTGTCTCCCTAAAATGTTATAAAGTAAACCGGTGACCTGACCACCTCAGGCA	1281
DB	1635	TTGTAACTCCTGTCTCCCTAAAATGTTATAAAGTAAACCGGTGACCTGACCACCTCAGGCA	1694
QY	1282	CACTTTCTCAGGACCTCTGAGAGTGATCCAGGCCATCGTAAAGTCANUTTGCTCAGA	1341
DB	1695	CACTTTCTCAGGACCTCTGAGAGTGATCCAGGCCATCGTAAAGTCANUTTGCTCAGA	1754

QY	1342	ATCAACCTCTTTAAATAATTTTACAGAAATTCGGGTTTTGGTTTACCAATAAGTCTCCACAAA	1401
DB	1755	ATCAACCTCTTTAAATAATTTTACAGAAATTTGGGTTTTGGTTTACCAATAAGTCTCCACAAA	1814
QY	1402	TATATGTCGAGAATCTTCAATTCGAAGCCTGCTCACCAAAATTTCAAAATGCCAACATCTC	1461
DB	1815	TATATGTCGAGAATCTTCAATTCGAAGCCTGCTCACCAAAATTTCAAAATGCCAACATCTC	1874
QY	1462	CCCATCCAATPACCTATTTCATCTTTGAGGTGAATCTPACFCAATAAATCTGTGAAGACC	1521
DB	1875	CCCATCCAATPACCTATTTCATCTTTGAGGTGAATCTPACFCAATAAATCTGTGAAGACC	1934
QY	1522	AGTGACGAGACCCCTTGTGCTAACCTGACATTTACTTCAATTTTCTTTTCTATGACTGG	1581
DB	1935	AGTGACGAGACCCCTTGTGCTAACCTGACATTTACTTCAATTTTCTTTTCTATGACTGG	1994
QY	1582	ATATTTTTCGATATAAACTTCGAGTAATAGTTCAAAAATTTAATAGTTTTGACATTCGCT	1641
DB	1995	ATATTTTTCGATATAAACTTCGAGTAATAGTTCAAAAATTTAATAGTTTTGACATTCGCT	2054
QY	1642	TTTCTGAGAAGAGAAATGAAAGTGTCAAAAAATAAAAAAGATGAAATGAAGCATATAT	1701
DB	2055	TTTCTGAGAAGAGAAATGAAAGTGTCAAAAAATAAAAAAGATGAAATGAAGCATATAT	2114
QY	1702	AATGTCAATTTTTTCAATTTTCTAGTCAACAGAGAATCGAAGATTTCTGTCAAATATT	1761
DB	2115	AATGTCAATTTTTTCAATTTTCTAGTCAACAGAGAATCGAAGATTTCTGTCAAATATT	2174
QY	1762	AGTAAAAATGAAAATAAACTTTGCTTATATTTTGTTCGACACACTAGTTAATTTAA	1821
DB	2175	AGTAAAAATGAAAATAAACTTTGCTTATATTTTGTTCGACACACTAGTTAATTTAA	2234
QY	1822	CCTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTTTAAAAATCGAAGC	1881
DB	2235	CCTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTTTAAAAATCGAAGC	2294
QY	1882	AAACTGGAAAAATATCCATCTAATTATGCTTTCTTTCCCAAGAAGTTTTTTAATGATG	1941
DB	2295	AAACTGGAAAAATATCCATCTAATTATGCTTTCTTTCCCAAGAAGTTTTTTAATGATG	2354
QY	1942	CCAGCTTCCTAATTTGGAGACAAAAGCCTTAATTGCAATGCAATTCAATATATTTTTT	2001
DB	2355	CCAGCTTCCTAATTTGGAGACAAAAGCCTTAATTGCAATGCAATTCAATATATTTTTT	2414
QY	2002	TGTATAGTTACAGTATACGAGTTTCAGTATCCCTTTAGATGAGATGCTTTGGACAGAGTG	2061
DB	2415	TGTATAGTTACAGTATACGAGTTTCAGTATCCCTTTAGATGAGATGCTTTGGACAGAGTG	2474
QY	2062	TTTTGGATTTTCAGATTTATTTTTGGATTTTGGATTTTCCATACATATAATGAGAGAGT	2121
DB	2475	TTTTGGATTTTCAGATTTATTTTTGGATTTTGGATTTTCCATACATATAATGAGAGAGT	2534
QY	2122	TGGAAATGGGATTCAGTCTAATCATATAAAATTCACCTATGTTTGATATACACCTTATCT	2181
DB	2535	TGGAAATGGGATTCAGTCTAATCATATAAAATTCACCTATGTTTGATATACACCTTATCT	2594
QY	2182	GAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAATATTTTATGCTGGAACAGAGTT	2241
DB	2595	GAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAATATTTTATGCTGGAACAGAGTT	2654
QY	2242	TGCGCAATTTGGACCATCAGAAAGCAGAGTGTCACTATTTCAAGTCAGTGTCAAAAAG	2301
DB	2655	TGCGCAATTTGGACCATCAGAAAGCAGAGTGTCACTATTTCAAGTCAGTGTCAAAAAG	2714
QY	2302	TTTCAGATGTTAAGCTGGTGATGTCAGTTGATGCCAGTGATCCGAGTACTTTGGGAAGCCA	2361
DB	2715	TTTCAGATGTTAAGCTGGTGATGTCAGTTGATGCCAGTGATCCGAGTACTTTGGGAAGCCA	2774
QY	2362	AGACAGTGGATCTCTTTGAGCCAGAGTTTGGGCCAGACTGCAAAACAGATGAGACC	2421
DB	2775	AGACAGTGGATCTCTTTGAGCCAGAGTTTGGGCCAGACTGCAAAACAGATGAGACC	2834
QY	2422	TCGTTTCTCAAAATAAATTTAAAAAATTTAGCCAGGTGGTGGTGGCACTGTAGTCCAG	2481

Db	2414	TGGAATAAGGGATTCAAGTCTAAATCAATAAAAA'TTCACTTAATGTTTGATATACACCTTATCT	2473
Qy	2182	GAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAATAATATTTATGCTCTGAACACAGATT	2241
Db	2474	GAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAATAATATTTATGCTCTGAACACAGATT	2533
Qy	2242	TGGCCACATTTGGACCATCAGAAAGCAGAAAGTGTCACATTTTCAAGTCAGTGCTCAAAAAG	2301
Db	2534	TGCGCACATTTGGACCATCAGAAAGCAGAAAGTGTCACATTTTCAAGTCAGTGCTCAAAAAG	2593
Qy	2302	TTTCAGATGTTAAGCTGGTGATCCAGTTTCATGCCAGTGATCCGAGTACTTTTCGGAGAGCCA	2361
Db	2594	TTTCAGATGTTAAGCTGGTGATCCAGTTTCATGCCAGTGATCCGAGTACTTTTCGGAGAGCCA	2653
Qy	2362	AGACAGGTGGATCTCTTTAGCCACGAGGTTTGAAGCCAGACTGCACAAACAGTGAAGACC	2421
Db	2654	AGACAGGTGGATCTCTTTAGCCACGAGGTTTGAAGCCAGACTGCACAAACAGTGAAGACC	2713
Qy	2422	TCGTTTCTCAAAATAATTTAAAAAATTAGCCAGGTGTGGTGGTGACACCTGTAGTCCCGAG	2488
Db	2714	TCGTTTCTCAAAATAATTTAAAAAATTAGCCAGGTGTGGTGGTGACACCTGTAGTCCCGAG	2773
Qy	2482	GTACTCAGGAGGCTGAGGTAGTAGGATTTGTTTCAGACTGGGAGGTTTGAAGCTGAACTGAG	2541
Db	2774	GTACTCAGGAGGCTGAGGTAGTAGGATTTGTTTCAGACTGGGAGGTTTGAAGCTGAACTGAG	2833
Qy	2542	CCAGGATCTTGCCACACACATTTCCAGCTTTGGGCAACAGAGTGAGACCTGTCTTC-AAAAAA	2600
Db	2834	CCAGGATCTTGCCACACACATTTCCAGCTTTGGGCAACAGAGTGAGACCTGTCTTC-AAAAAA	2893
Qy	2601	AAAAAAAAGTTTTCAGATTTTTGGAGCATTTTCAGATCTTTCAGATTAGGATTTTTCAACCTGT	2660
Db	2894	AAAAAAAAGTTTTCAGATTTTTGGAGCATTTTCAGATCTTTCAGATTAGGATTTTTCAACCTGT	2953
Qy	2661	ACTGACCTTTTGTTCATGTGACAGACATTAATCAATAGGTGGACTCCAGATTAATCTCATTTG	2720
Db	2954	ACTGACCTTTTGTTCATGTGACAGACATTAATCAATAGGTGGACTCCAGATTAATCTCATTTG	3013
Qy	2721	CTGTATACACATTTTGCTCTCTATTCCACGAATTTCTTATGCCCTCTTGTGGTGATTTTA	2780
Db	3014	CTGTATACACATTTTGCTCTCTATTCCACGAATTTCTTATGCCCTCTTGTGGTGATTTTA	3073
Qy	2781	ATGTGCGGAAGGGAAAAAACAATAGAAATTTTGCAAATTTCTAGAAAAAGTCATTTCTGCAAAAATA	2840
Db	3074	ATGTGCGGAAGGGAAAAAACAATAGAAATTTTGCAAATTTCTAGAAAAAGTCATTTCTGCAAAAATA	3133
Qy	2841	TGTGAGTCTGTAGATATTAGCCAAATTTTAGGAAAAATTCAGAAAAATTTTACTTTTCGTGTC	2900
Db	3134	TGTGAGTCTGTAGATATTAGCCAAATTTTAGGAAAAATTTTACTTTTCGTGTCGTGTC	3193
Qy	2901	TGCCCTTTGTAGCTGTTTTATCATATAATACCTTATTTTGTAAATAAAATTAATTTTAATTT	2960
Db	3194	TGCCCTTTGTAGCTGTTTTATCATATAATACCTTATTTTGTAAATAAAATTAATTTTAATTT	3253
Qy	2961	GAGTAAACAATCTGGAAATTTATCAGAGAAGGGCAAGCAATAGGTTTAATAACAGTATTGAT	3020
Db	3254	GAGTAAACAATCTGGAAATTTATCAGAGAAGGGCAAGCAATAGGTTTAATAACAGTATTGAT	3313
Qy	3021	TGGTAGAAGGAAAGCTTGAATTCGAAGAGCATCAATGTCTTCTGTGTGTTTACCCATAAGCC	3080
Db	3314	TGGTAGAAGGAAAGCTTGAATTCGAAGAGCATCAATGTCTTCTGTGTGTTTACCCATAAGCC	3373
Qy	3081	ACAGCAGATGTCCTAACTCTTTCCGAGATCTAGTTTTTTTTCAGCAAAAGCAGGATTTTAAAGAAAT	3140
Db	3374	ACAGCAGATGTCCTAACTCTTTCCGAGATCTAGTTTTTTTTCAGCAAAAGCAGGATTTTAAAGAAAT	3433
Qy	3141	GTAACTACTTATGTGGTTATGAAGAAACAATAGAAATCATTTGCTATAGTGCTTTTTTAA	3200
Db	3434	GTAACTACTTATGTGGTTATGAAGAAACAATAGAAATCATTTGCTATAGTGCTTTTTTAA	3493
Qy	3201	CCTGTAAATTTTGTGAAGCTTATCTTTTATGCATATAAAATTTTGAACATTTTACATTTGT	3260
Db	3494	CCTGTAAATTTTGTGAAGCTTATCTTTTATGCATATAAAATTTTGAACATTTTACATTTGT	3553

Qy	3261	TTATATTTTAAATCAGTCTTTTACTCAAGTGTGATATATACAGAGAAATGTAAACCACTGTA	3322
Db	3554	TTATATTTTAAATCAGTCTTTTACTCAAGTGTGATATATACAGAGAAATGTAAACCACTGTA	3613
Qy	3321	AGGGTACAGTTATAAGAAATTTTGTGCAAACTGATTTTCAACCATGTAGTCACCTCCTTAATGAA	3380
Db	3614	AGGGTACAGTTATAAGAAATTTTGTGCAAACTGATTTTCAACCATGTAGTCACCTCCTTAATGAA	3673
Qy	3381	GAGACAGAACACGTACATCTCCCCAGAGAGTTCCACAGATGCTCCTTTTCCCTGAGTTTCA	3440
Db	3674	GAGACAGAACACGTACATCTCCCCAGAGAGTTCCACAGATGCTCCTTTTCCCTGAGTTTCA	3733
Qy	3441	CCAGTCTCGCAACCAATGATCTGCTTCGTATTAATTAATAACTGTTCTAGATATTTTCTGAGC	3500
Db	3734	CCAGTCTCGCAACCAATGATCTGCTTCGTATTAATTAATAACTGTTCTAGATATTTTCTGAGC	3793
Qy	3501	AATGTACCTTTTCCATATTTTATTTTGTGTGTGAAGGCTTCTTTTAACTCATTAATAATTT	3560
Db	3794	AATGTACCTTTTCCATATTTTATTTTGTGTGTGAAGGCTTCTTTTAACTCATTAATAATTT	3853
Qy	3561	TTTGAGATTCATCTPATGTTTAAATGTTCTATCAGTAGTTGTACATCTTTTCTGTCTCAGCA	3620
Db	3854	TTTGAGATTCATCTPATGTTTAAATGTTCTATCAGTAGTTGTACATCTTTTCTGTCTCAGCA	3913
Qy	3621	TATCACCATATAGATATACATATAATTTTGTGTTAACTTAATCACTCATGGATATGTAGGATAT	3680
Db	3914	TATCACCATATAGATATACATATAATTTTGTGTTAACTTAATCACTCATGGATATGTAGGATAT	3973
Qy	3681	TTAAAGTTTTTGACATTAATGAATAAAGTGGCTATAAATGAAAAAAGAAAAA 3735	
Db	3974	TTAAAGTTTTTGACATTAATGAATAAAGTGGCTATAAATGAAAAAAGAAAAA 4028	
RESULT 3			
LOCUS	CQ841482	4136 bp	DNA linear PAT 02-AUG-2004
DEFINITION	Sequence 129 from Patent EP1440981.		
ACCESSION	CQ841482		
VERSION	CQ841482.1	GI:50893269	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Esakaya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yanamoto, J., Isono, Y., Nagai, K. and Irie, R.		
JOURNAL	Full-length human cdna		
FEATURES	Patent: EP 1440981-A 129 28-JUL-2004;		
source	Research Association for Biotechnology (JP)		
	Location/Qualifiers		
	1..4136		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	95.4%;	Score 3590;	DB 6; Length 4136;
Best Local Similarity	99.3%;	Pred. No. 0;	
Matches 3617;	Conservative	0; Mismatches	20; Indels 5; Gaps 1;
Qy	82	AGAAAGGAACCCAGATGAGGAAGAAGAAACGGCAGGAAATGACTACATTTTGAAGAAA	141
Db	495	AGATGGGAAATATCTCTCTCGAAACTCAGGAGGAAATGACTACATTTTGAAGAAA	554
Qy	142	CTTCAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTTAATCAGAAAAACGAG	201
Db	555	CTTCAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTTAATCAGAAAAACGAG	614
Qy	202	AATGCCAGTGGTTCTGAAGAAAGTGTGTACACTGTCAATTAATCAATCCCCATCAGAGA	261
Db	615	AATGCCAGTGGTTCTGAAGAAAGTGTGTACACTGTCAATTAATCAATCCCCATCAGAGA	674



```
QY 2281 TTCAAGTCAGTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTCATGCCAGTGA 2340
Db 2281 TTCAAGTCAGTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTCATGCCAGTGA 2340
QY 2341 TCCGAGTACTTTGGGAAGCCCAAGACAGTGGATCTCTTGAGCCCAAGGAGTTTGAGGCCAG 2400
Db 2341 TCCGAGTACTTTGGGAAGCCCAAGACAGTGGATCTCTTGAGCCCAAGGAGTTTGAGGCCAG 2400
QY 2401 ACTGCAACAACAGTGGAGACCTCGTTTCTACAAAATAATTAATAAATTTAGCCAGGTGGT 2460
Db 2401 ACTGCAACAACAGTGGAGACCTCGTTTCTACAAAATAATTAATAAATTTAGCCAGGTGGT 2460
QY 2461 GGTGCACACCTGTAGTCCAGGTACTCAGAGAGCTGAGTAGTAGGATTTGTTGAGACTG 2520
Db 2461 GGTGCACACCTGTAGTCCAGGTACTCAGAGAGCTGAGTAGTAGGATTTGTTGAGACTG 2520
QY 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCAACATTCAGAGCTTGGCCACAGAG 2580
Db 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCAACATTCAGAGCTTGGCCACAGAG 2580
QY 2581 TGAGACCTGTCTCAAAAAAAGTTTCAGATTTTGAGAGCATTTTCAGATCTTCAG 2640
Db 2581 TGAGACCTGTCTCAAAAAAAGTTTCAGATTTTGAGAGCATTTTCAGATCTTCAG 2640
QY 2641 ATTAGGATTTTCAACCTGTACTGACCTTTTAGTCATTGACAGCATTAATCAATAGGTG 2700
Db 2641 ATTAGGATTTTCAACCTGTACTGACCTTTTAGTCATTGACAGCATTAATCAATAGGTG 2700
QY 2701 GACTCCAGATAACTCAITTTGCTGTATACACATTTTGCCTCTCTATTCAACGAATTCCTAT 2760
Db 2701 GACTCCAGATAACTCAITTTGCTGTATACACATTTTGCCTCTCTATTCAACGAATTCCTAT 2760
QY 2761 GCCCTCTTGTGTTGTTTAAATGTCGGAAGGAAACAAATAGAAATTTTGCATTTCTAGA 2820
Db 2761 GCCCTCTTGTGTTGTTTAAATGTCGGAAGGAAACAAATAGAAATTTTGCATTTCTAGA 2820
QY 2821 AAAGTCATTCTGTCAAAATATGTCAGTCTGAGATATTAGCCAAATTTTAGGAAATGAC 2880
Db 2821 AAAGTCATTCTGTCAAAATATGTCAGTCTGAGATATTAGCCAAATTTTAGGAAATGAC 2880
QY 2881 AAAATTTTACTTTTCGTCCTCTGTTAGTCTGTTTATGATAPAAATACCTTATTTGT 2940
Db 2881 AAAATTTTACTTTTCGTCCTCTGTTAGTCTGTTTATGATAPAAATACCTTATTTGT 2940
QY 2941 AATAAAATTAATTTTAAATTTGAGTAAACATCTGGAATATTCAGAGAAGGGCAACAATA 3000
Db 2941 AATAAAATTAATTTTAAATTTGAGTAAACATCTGGAATATTCAGAGAAGGGCAACAATA 3000
QY 3001 GGTAAATAAACAGTATTGATTGGTAGAAGGAACGTTGAAATCCAGAGCATCAATGTCTT 3060
Db 3001 GGTAAATAAACAGTATTGATTGGTAGAAGGAACGTTGAAATCCAGAGCATCAATGTCTT 3060
QY 3061 CTGGTGGTTTCAACATAAGCCACAGAGATGCTTAAATCTTTCCGAGATCTAGTTTTCAG 3120
Db 3061 CTGGTGGTTTCAACATAAGCCACAGAGATGCTTAAATCTTTCCGAGATCTAGTTTTCAG 3120
QY 3121 CAAGCAGATTTAGAAATGTAATCTTATGTTGTTATGATGATGATGATGATGATGATGAT 3180
Db 3121 CAAGCAGATTTAGAAATGTAATCTTATGTTGTTATGATGATGATGATGATGATGATGAT 3180
QY 3181 GCTGTATAAGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGATATAAAT 3240
Db 3181 GCTGTATAAGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGATATAAAT 3240
QY 3241 ATTTGAACATTTTACATGTTTATATTTTAAATCACTTTTACTCAAGTGTGATATATAC 3300
Db 3241 ATTTGAACATTTTACATGTTTATATTTTAAATCACTTTTACTCAAGTGTGATATATAC 3300
QY 3301 AAGAAATGTAACCTGTAAGGGTAGGTTATAGAAATTTTGTCAAAATGTTTACCCCA 3360
Db 3301 AAGAAATGTAACCTGTAAGGGTAGGTTATAGAAATTTTGTCAAAATGTTTACCCCA 3360
QY 3361 TGTAGTCACCTCTTATGAAGAGACAGACAGTATACATCTCTCCCAAGAAAGTTCCACAGTG 3420
```

```
Db 3361 TGTAGTCACCTCTTATGAAGAGACAGACAGTATACATCTCTCCCAAGAAAGTTCCACAGTG 3420
QY 3421 CTCCTTTTCCCTGAGTTTCACCACTCTGGCAACCAATGATCTCTCGTATATAATATAA 3480
Db 3421 CTCCTTTTCCCTGAGTTTCACCACTCTGGCAACCAATGATCTCTCGTATATAATATAA 3480
QY 3481 CTGTTCTAGATATTGTTAGCAATGTACCCCTTTTCCATATTTATTTTGTGTGTAGGCTT 3540
Db 3481 CTGTTCTAGATATTGTTAGCAATGTACCCCTTTTCCATATTTATTTTGTGTGTAGGCTT 3540
QY 3541 CTTTGTAGTCATTATAATATTTTTCAGATTCATCTATGTTTAAATGTTCTATCAGTAGTGT 3600
Db 3541 CTTTGTAGTCATTATAATATTTTTCAGATTCATCTATGTTTAAATGTTCTATCAGTAGTGT 3600
QY 3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTTCTTAAATCTTAATCA 3660
Db 3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTTCTTAAATCTTAATCA 3660
QY 3661 CTGATGGATATGTTAGGATATTAAAGTTTTCACATTTATGAATAAGTGGCTATATAATGAA 3720
Db 3661 CTGATGGATATGTTAGGATATTAAAGTTTTCACATTTATGAATAAGTGGCTATATAATGAA 3720
QY 3721 AAAAAAAGGCTTCTCAGCATATCACCATATAGATATATCTTCTTAAATCTTAATCA 3762
Db 3721 AAAAAAAGGCTTCTCAGCATATCACCATATAGATATATCTTCTTAAATCTTAATCA 3762
```

## RESULT 2

AK126682  
LOCUS Homo sapiens cDNA FLJ44728 fis, clone BRACE3024537. 4032 bp mRNA linear PRI 19-FEB-2004

## DEFINITION

AK126682

## ACCESSION

AK126682.1 GI:34533254

## VERSION

oligo capping; fis (full insert sequence).

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1

## AUTHORS

Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,

Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,

Kanihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,

Yanamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,

Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,

Kikuchi, H., Murakawa, K., Kawakami, B., Suzuki, Y., Sugano, S.,

Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,

Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

NEBO human cDNA sequencing project

Unpublished

2 (bases 1 to 4032)

## JOURNAL

Isogai, T. and Yamamoto, J.

## REFERENCE

Direct Submission

## AUTHORS

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEBO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- &amp; 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: Reverse Proteomics Research Institute, HRI and

RAB.

## FEATURES

Location/Qualifiers

1..4032

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="BRACE3024537"

/tissue\_type="cerebellum"

Db 61 GGAGAGAAATCAAAAGAACCCAGAAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAA 120  
Qy 121 ATGACTACATTTGAAAGAAAATTTCAAGATCAAGATAGAAAGAACCCAGAGAGTTTTCATCC 180  
Db 121 ATGACTACATTTGAAAGAAAATTTCAAGATCAAGATAGAAAGAACCCAGAGAGTTTTCATCC 180  
Qy 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTTCTGAAGAGAGTGTGCTACACTGTCAAT 240  
Db 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTTCTGAAGAGAGTGTGCTACACTGTCAAT 240  
Qy 241 AATCACATCCCCCATCAGAGATCTCCCTGAGCTCCCAATGATGATGCTATGAGAACAT 300  
Db 241 AATCACATCCCCCATCAGAGATCTCCCTGAGCTCCCAATGATGATGCTATGAGAACAT 300  
Qy 301 GACTCCCTCACAAGAAAGTGAGACAGTTTAGAAGAAAGGTGAGAGACAGAAATATGCCCC 360  
Db 301 GACTCCCTCACAAGAAAGTGAGACAGTTTAGAAGAAAGGTGAGAGACAGAAATATGCCCC 360  
Qy 361 CTTAGGACTCTGTTAGTAGGCTTGTTCCTGACCCCATGAGCATGATTAAGAAGTTGTG 420  
Db 361 CTTAGGACTCTGTTAGTAGGCTTGTTCCTGACCCCATGAGCATGATTAAGAAGTTGTG 420  
Qy 421 TTTCACACATAAATCTCAAGCTGCTTTATCACTTTCCAGCAATGAAGACAAATGCAGAA 480  
Db 421 TTTCACACATAAATCTCAAGCTGCTTTATCACCTTCCAGCAATGAAGACAAATGCAGAA 480  
Qy 481 TAGCAGACTCTGGGAAAGTTGTTCAACCTGAGCAGTGCATGAACATTCCTTTCTGGCTA 540  
Db 481 TAGCAGACTCTGGGAAAGTTGTTCAACCTGAGCAGTGCATGAACATTCCTTTCTGGCTA 540  
Qy 541 AAGTTTAGAATAATATCTATTATATATATCTTAGGCAACTCTGATATGTTGGCATCTG 600  
Db 541 AAGTTTAGAATAATATCTATTATATATCTTAGGCAACTCTGATATGTTGGCATCTG 600  
Qy 601 TGGCTTAGGTAATCATAGAAATTTGACAAATTTGACAAATTTGACAAATTTGACAAATTTG 660  
Db 601 TGGCTTAGGTAATCATAGAAATTTGACAAATTTGACAAATTTGACAAATTTGACAAATTTG 660  
Qy 661 TTGTAAGTTTAGGACATGAGAGTGATATAAATAAATTTCTTAGGCAATAATGTAATAA 720  
Db 661 TTGTAAGTTTAGGACATGAGAGTGATATAAATAAATTTCTTAGGCAATAATGTAATAA 720  
Qy 721 TGAATAATAATTTCTAATCCCCCTGACTAATCTGAATGAGCCCTCTCTAGGCCAAGAGA 780  
Db 721 TGAATAATAATTTCTAATCCCCCTGACTAATCTGAATGAGCCCTCTCTAGGCCAAGAGA 780  
Qy 781 CTTGAGATGAACCTGAAAGACTGAAATTTCTGGCCATGATAGGAGGAGGTGAGACACACC 840  
Db 781 CTTGAGATGAACCTGAAAGACTGAAATTTCTGGCCATGATAGGAGGAGGTGAGACACACC 840  
Qy 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTGACAGGATGAGTCAATAGACT 900  
Db 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTGACAGGATGAGTCAATAGACT 900  
Qy 901 GATGAATAAGACTGATTTGGCAATAGAGTCCCAATTTCCAACTGACTCTGTTGTAGAT 960  
Db 901 GATGAATAAGACTGATTTGGCAATAGAGTCCCAATTTCCAACTGACTCTGTTGTAGAT 960  
Qy 961 CACACACTGTCTGAGGATTTCCATCTATGAGACTTTGCTTACATACAGAGACCTTGGTT 1020  
Db 961 CACACACTGTCTGAGGATTTCCATCTATGAGACTTTGCTTACATACAGAGACCTTGGTT 1020  
Qy 1021 TCCACAACCCCTTTATTTAGCTTAAGCAATTCAGACAACTTTGAATCTACCTATGACCT 1080  
Db 1021 TCCACAACCCCTTTATTTAGCTTAAGCAATTCAGACAACTTTGAATCTACCTATGACCT 1080  
Qy 1081 CAAAGCTTAATCTTTTCAACCAATTTGCCAATCAGACAACTTTGAATCTACCTATGACCT 1140  
Db 1081 CAAAGCTTAATCTTTTCAACCAATTTGCCAATCAGACAACTTTGAATCTACCTATGACCT 1140  
Qy 1141 GTAGACTCTCTCTGCTTCAAGATCTTGGCTCTTTTAAAGCTGAACCGATGTGCACTTTCCA 1200  
Db 1141 GTAGACTCTCTCTGCTTCAAGATCTTGGCTCTTTTAAAGCTGAACCGATGTGCACTTTCCA 1200

Qy 1201 TTTAATGATTTATGTCTTTTGTAACTCTCTCTCCCTAAATATGTAATAAGTAAACG 1260  
Db 1201 TTTAATGATTTATGTCTTTTGTAACTCTCTCTCCCTAAATATGTAATAAGTAAACG 1260  
Qy 1261 GTGACCTGACCACTCTGAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320  
Db 1261 GTGACCTGACCACTCTGAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320  
Qy 1321 GGTAAAGTCATGTTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380  
Db 1321 GGTAAAGTCATGTTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380  
Qy 1381 TTACCAATAAGTCTCCCAAAATATATGTCCCAAGAACTCTTCAATTTCCAAGCCTGCTCACCA 1440  
Db 1381 TTACCAATAAGTCTCCCAAAATATATGTCCCAAGAACTCTTCAATTTCCAAGCCTGCTCACCA 1440  
Qy 1441 AATTTCAAAATGCCAAATCTCCCAATCTCCCAATCTCCCAATCTCCCAATCTCCCAATCTA 1500  
Db 1441 AATTTCAAAATGCCAAATCTCCCAATCTCCCAATCTCCCAATCTCCCAATCTCCCAATCTA 1500  
Qy 1501 CTCAATAAACTGTGTAAAGCAGTGAACCACTGACAGCCCTTTGCTTAACCTGACATTTTCAAT 1560  
Db 1501 CTCAATAAACTGTGTAAAGCAGTGAACCACTGACAGCCCTTTGCTTAACCTGACATTTTCAAT 1560  
Qy 1561 TTTTCTTTTCTATGTAAGTATTTTTCATATATAAATTTGAGTGTATTTTCAATTTTCAAT 1620  
Db 1561 TTTTCTTTTCTATGTAAGTATTTTTCATATATAAATTTGAGTGTATTTTCAATTTTCAAT 1620  
Qy 1621 TAATAGTTTGTGACATTTGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680  
Db 1621 TAATAGTTTGTGACATTTGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680  
Qy 1681 AAGATGAATGAAGCATATATATTTCTCAATTTTCTCAATTTTCTAGTCAACAGAGATC 1740  
Db 1681 AAGATGAATGAAGCATATATATTTCTCAATTTTCTCAATTTTCTAGTCAACAGAGATC 1740  
Qy 1741 GAAGGATTTCTGTTTCAAAATTTAGTAAAAATTTGAAAAATTTGAAAGTGTCAAAAATAAAAA 1800  
Db 1741 GAAGGATTTCTGTTTCAAAATTTAGTAAAAATTTGAAAAATTTGAAAGTGTCAAAAATAAAAA 1800  
Qy 1801 GCAACACACTAGTAAATTTAACTGAGTGTAGTATCTCTACCGAAGGTGGATGTAGT 1860  
Db 1801 GCAACACACTAGTAAATTTAACTGAGTGTAGTATCTCTACCGAAGGTGGATGTAGT 1860  
Qy 1861 TTCTGCTTTTAAAAATTTCAAGCAAACTGAAAAATTTGAAATTTGCTTTCTTTTCCC 1920  
Db 1861 TTCTGCTTTTAAAAATTTCAAGCAAACTGAAAAATTTGAAATTTGCTTTCTTTTCCC 1920  
Qy 1921 AAGAAGTTTAAATGAATGCGAGCTTCTTAATTTGGAGACAAAAAGCCTTAAATGACAA 1980  
Db 1921 AAGAAGTTTAAATGAATGCGAGCTTCTTAATTTGGAGACAAAAAGCCTTAAATGACAA 1980  
Qy 1981 TGCAATTCATTAATATTTTGTAGTATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Db 1981 TGCAATTCATTAATATTTTGTAGTATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Qy 2041 AGATGCTTGGGACCAAGAGTGTGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTT 2100  
Db 2041 AGATGCTTGGGACCAAGAGTGTGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTT 2100  
Qy 2101 CCATACATATAATGAGAGAGTTGGAAATTTGGAATTTCAAGTCTAAATCATAAAAATTCACCTA 2160  
Db 2101 CCATACATATAATGAGAGAGTTGGAAATTTGGAATTTGGAATTTCAAGTCTAAATCATAAAAATTCACCTA 2160  
Qy 2161 TGTGATATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATTTTAAATTA 2220  
Db 2161 TGTGATATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATTTTAAATTA 2220  
Qy 2221 TTTTATGCTGAAAACAGAGTGTTCGCGACATTTGGACCATCGAAGAGTGTCACTAT 2280  
Db 2221 TTTTATGCTGAAAACAGAGTGTTCGCGACATTTGGACCATCGAAGAGTGTCACTAT 2280

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 18, 2004, 12:52:13 : Search time 15375 Seconds  
(without alignments)  
11570.982 Million cell updates/sec

Title: US-10-005-907-1  
Perfect score: 3762  
Sequence: 1 gagaaacgagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3762	100.0	3762	6	AX505122	AX505122 Sequence
2	3596.6	95.6	4032	6	AK126682	AK126682 Homo sapi
3	3590	95.4	4136	6	CQ841482	CQ841482 Sequence
4	3590	95.4	4136	6	AK124520	AK124520 Homo sapi
5	3522.2	93.6	185467	9	AL606804	AL606804 Human DNA
6	3511.6	93.3	141268	2	AC074365	AC074365 Homo sapi
7	2604.8	69.2	2648	9	BC024174	BC024174 Homo sapi
8	1673.4	44.5	1709	6	CQ842314	CQ842314 Sequence
9	1673.4	44.5	1709	9	AK123798	AK123798 Homo sapi
10	1504.6	40.0	1924	9	AB060908	AB060908 Macaca fa
11	1068	28.4	1128	6	CQ488422	CQ488422 Sequence
12	1068	28.4	1128	6	CQ488563	CQ488563 Sequence
13	1068	28.4	1128	6	CQ488741	CQ488741 Sequence
14	1068	28.4	1128	6	CQ488930	CQ488930 Sequence
15	1068	28.4	1128	6	CQ494252	CQ494252 Sequence
16	1068	28.4	1128	6	CQ494397	CQ494397 Sequence
17	1068	28.4	1128	6	CQ494580	CQ494580 Sequence
18	1068	28.4	1128	6	CQ494775	CQ494775 Sequence
19	994.8	26.4	121353	2	AC023386	AC023386 Homo sapi

C	20	697.8	18.5	121353	2	AC023386	AC023386 Homo sapi
	21	394.2	10.5	433	6	CQ469395	CQ469395 Sequence
	22	393.8	10.5	397	6	CQ478563	CQ478563 Sequence
	23	392.2	10.4	421	6	CQ499427	CQ499427 Sequence
	24	392.2	10.4	437	6	CQ499760	CQ499760 Sequence
	25	392.2	10.4	437	6	CQ508397	CQ508397 Sequence
	26	392.2	10.4	437	6	CQ508729	CQ508729 Sequence
	27	391.8	10.4	404	6	CQ478228	CQ478228 Sequence
	28	375	10.0	446	6	CQ469060	CQ469060 Sequence
	29	354	9.4	70311	9	AC002390	AC002390 Human DNA
	30	353.6	9.4	174875	2	AC024682	AC024682 Homo sapi
	31	342.8	9.1	359	6	CQ477717	CQ477717 Sequence
	32	342.2	9.1	400	6	CQ468549	CQ468549 Sequence
	33	342.2	9.1	401	6	CQ498920	CQ498920 Sequence
	34	342.2	9.1	401	6	CQ507889	CQ507889 Sequence
	35	342.2	9.1	401	6	CQ508110	CQ508110 Sequence
	36	340.6	9.1	149490	9	AL355344	AL355344 Human DNA
	37	340.6	9.1	168312	2	AC007484	AC007484 Homo sapi
	38	340.6	9.1	191161	2	AC016043	AC016043 Homo sapi
	39	300.8	8.0	150424	9	AL391811	AL391811 Human DNA
	40	300.8	8.0	190793	2	AC034298	AC034298 Homo sapi
	41	300.8	8.0	199594	9	AC093560	AC093560 Homo sapi
	42	295.8	7.9	168055	9	AC090796	AC090796 Homo sapi
	43	295.8	7.9	191037	9	AC091561	AC091561 Homo sapi
	44	295.8	7.9	191234	2	AC018681	AC018681 Homo sapi
	45	277	7.4	98697	9	AC004854	AC004854 Homo sapi

ALIGNMENTS

RESULT 1  
AX505122  
LOCUS AX505122 3762 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 1 from Patent WO0246389.  
ACCESSION AX505122  
VERSION AX505122.1 GI:23386429  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Nocka,K., Pirozzi,G. and Einstein,R.  
TITLE Novel genes associated with allergic hypersensitivity and mast cell activation  
JOURNAL Patent: WO 0246389-A 1 13-JUN-2002;  
UCB, S.A. (BE)  
FEATURES  
source Location/Qualifiers  
1..3762  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
25..432  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD48808.1"  
/db\_xref="GI:23386430"  
/translation="MGNYLLRLKSLCLGENQKPKXGNPDEERKROEMTTFERKLODQD  
KKQSVSTNOENENGSGSEVCVTVINHIFHQSSLSNDDGYNIDSLTRKVRQF  
RRSETEYALLTSVSRPCSTHEHDYEVVFFH"

Query Match 100.0%; Score 3762; DB 6; Length 3762;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAGAAACCGAGTCACGTGTGAAAGATGGGAAATATCTCTCGGAAAACCTCAGTTCCTG 60  
Db 1 GAGAAACCGAGTCACGTGTGAAAGATGGGAAATATCTCTCGGAAAACCTCAGTTCCTG 60  
QY 61 GGAGAGAAATCAAGAGAGCCAGAGAAAGGAAACCCAGATCAGGAGAAACGCGAGGAA 120



**This Page Blank (uspto)**

```
; LENGTH: 130207
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(130207)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1372

Query Match          5.3%; Score 201.2; DB 13; Length 130207;
Best Local Similarity 72.2%; Pred. No. 1.7e-28;
Matches 322; Conservative 0; Mismatches 108; Indels 16; Gaps 4;

QY 954 TGTAGATCACACTGTCTGAGGAGTCCATCTATGAGACTTTGTCTTACATAACAGAGAC 1013
Db      |||||
QY 935 TTTACTGCCTATTCTCTGAAGGCTGATGCCTGTGAACCTTTGTCTTATACAGAAC 876
Db      |||||
QY 1014 CTTGGTTCCACACACC-CTTTATTTTACGTAAGCATCTTTTCTACTGACTCTTTAAG 1072
Db      |||||
QY 875 CTTGGTCTCTATAGCGTCCCTCATATTAACCTGAACATTTCTTTCTATTGACTTC--AAG 818
Db      |||||
QY 1073 TCTTTAGACAAAGCTTAACCTTTTCAACCAATGCGCAATCAGACAACTTTGAATCTACC 1132
Db      |||||
QY 817 TCTTTAGACACAGTTGACCTTTTAAACCAATGCGCAATCAGAAATCTTAGACTCCACC 758
Db      |||||
QY 1133 TATGACCTGTAACTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTC 1192
Db      |||||
QY 757 TAAGACCTGTAA--CCAGCCCTCCAAAGGTATCTTGTCTTTTAGCTGAATCAATGTAT 700
Db      |||||
QY 1193 ACTTTCCATTAATGATTTATGCTTTGCTGTAACTCTCTGCTCCCTAAATGTATAA 1252
Db      |||||
QY 699 ACCTTCCATGTATGTACCTTATGATTTTACCTAAATTCCTGCTCCCTAAATGTACAA 640
Db      |||||
QY 1253 AGTAAACGGTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGATCC 1312
Db      |||||
QY 639 ACCAATGTATACCTGACCGCTTGGGCACACTTTCTCAGGATCTCTTGAGACTGTCCCC 580
Db      |||||
QY 1313 CAGGCCATGTAAGT-----CATGTTGGCTCAGAAATCAACCTCTTTAAATATTT 1361
Db      |||||
QY 579 CAGGCCATGTCACCTCATATATGGCAAGCATATGGCTCAGAAATCAACCTCTTTAAATATTT 520
Db      |||||
QY 1362 TACAGAAATTTGGTTTGGTTACCAA 1387
Db      |||||
QY 519 TACAGAGCAGTAATTTGGGAGGCTAA 494
Db      |||||

RESULT 14
US-09-764-869-2326
; Sequence 2326, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2326
; LENGTH: 17239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2326

Query Match          5.3%; Score 200.4; DB 9; Length 17239;
Best Local Similarity 74.6%; Pred. No. 1e-28;
Matches 252; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 2287 TCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGATGCAGTTCATGCCAGTGCAG 2346
Db      |||||
QY 10164 TCCTTGATTAATAAAGTCAGATTCCTGAGGCTGAGTGTGTCTCACACCTGTAATCCAG 10223
Db      |||||
QY 2347 TACTTTGGGAAGCCAAAGACAGAGTGGATCTCTTTGAGCCAGGAGTTTGGGCCAGACTGCA 2406
Db      |||||
QY 10224 CACTTTGGGAGGCCAAGGCAGGTGGATCACTTGAGGTCAAGAGTTTGAGACCCGCTGGC 10283
Db      |||||
QY 2407 CAACACAGTGAACCTCGTTTCTACAAATAATTAATAAATAATAGCCAGGTGTGGTGGTGA 2466
Db      |||||
QY 10284 CAACATGGTGAACCCCTGTCTCTACTATAAATAATCAAAAATAGCCAGGTGTGGTGGTGA 10343
Db      |||||
QY 2467 CACCTGTAGTCCAGGCTACTCAGGAGGCTGAGGTAGTAGATTGTTTGAGACTGGGAGGT 2526
Db      |||||
QY 10344 TGCCTGTAGTCCAGTACTCGGAGGCTGAGGAGGAGAAATCGCTTTGAACCCAGGAGGT 10403
Db      |||||
QY 2527 TGAGGCTGAACCTGAGCCAGGATCTTGCCACCACTTCCAGCTTGGGCAACAGAGTCAGAC 2586
Db      |||||
QY 10404 GGAGGTTCAGTGAAGGAGATTCGCCATTGCATCCAGCTTGGGCAACAGAGTCAGAC 10463
Db      |||||
QY 2587 CTGTCTCAAAAAGTTTTCAGATTTTGAAGTTCAGATTTTGAAGTTCAGATTTTGAAGTTC 2624
Db      |||||
QY 10464 TCCGTCTCAAAAAGTTTTCAGATTTTGAAGTTCAGATTTTGAAGTTCAGATTTTGAAGTTC 10501
Db      |||||

Search completed: November 19, 2004, 01:44:09
Job time : 1747 secs
```

<div><div>Publication No. US20020182586A1</div><div>GENERAL INFORMATION: APPLICANT: Engelhard, David W. APPLICANT: Morris, Eric K. TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR TITLE OF INVENTION: CANCER FILE REFERENCE: 529452000122 CURRENT APPLICATION NUMBER: US/10/087,192 CURRENT FILING DATE: 2002-03-01 PRIOR FILING DATE: 2000-12-22 PRIOR FILING DATE: 2000-12-22 PRIOR FILING DATE: 2001-03-02 NUMBER OF SEQ ID NOS: 2059 SOFTWARE: FastSeq for Windows Version 4.0 SEQ ID NO 1384 LENGTH: 95960 TYPE: DNA ORGANISM: Homo sapiens FEATURE: NAME/KEY: misc feature LOCATION: (1)-(95960) OTHER INFORMATION: n = A,T,C or G US-10-087-192-1384</div></div> <div><div>Query Match 5.3%; Score 201.2; DB 13; Length 95960; Best Local Similarity 72.2%; Pred. No. 1.5e-28; Matches 322; Conservative 0; Mismatches 108; Indels 16; Gaps 4;</div><div><div>QY 954 TGTAGATCACACACTGCTGAGGATTCATCTATGAGACTTTGTCTACATACAGAGAC 1013 DB 4743 TTACTGCTCTATCTCTCTGAGGCTGATGCTGTGAACTTTGTCTTATACAGAAC 4802 QY 1014 CTGGTTTCCAAACC-CCTTTATTTAGCTAAAGCATTTCTTCTACTGACTTTCTTAAG 1072 DB 4803 CTGGTCTCTATAGCTCCCTCATATTAAGTGAACATTTCTTCTATTGACTTC--AAG 4860 QY 1073 TCTTTAGCAAGCTTAACTCTTTTCAACCAATGCGCAATCAGACAACTTTGAACTACC 1132 DB 4861 TCTTTAGACACAGTTGAACTCTTTTAAACCAATGCGCAATCAGAAATCTTAGACTCCACC 4920 QY 1133 TATGACTGTAGCTCTCTCTGCTTCAAGATCTTTGCTCTTTAAGCTGAACCGATGTGC 1192 DB 4921 TAAGACTGTAA--CCACCCCTCCAAGGTATCTTGTCTTTAGGCTGAATCAATGTAT 4978 QY 1193 ACTTTCCATTTAATGATTTATGCTTTTGTGTTAACTCTCTCCCTTAAATGTATAA 1252 DB 4979 ACCTTCATGTTGACTTATGATTTTACCTAAATTCCTGCTCCCTTAAATGTACAA 5038 QY 1253 AGTAAAGGTGACCTGACCACTCAGGCACACTTTCTCAGCACTCTCTGAGGTGTATCC 1312 DB 5039 ACCAATGTATACCTGACCGCTTTGGGCACACTTTCTCAGGATCTCTTGAGACTGTCCCC 5098 QY 1313 CAGGCCATGGTAAAGT-----CATGTTGGCTCAGAAATCAACCTCTTTAAATATTT 1361 DB 5099 CAGGCCATGGTCACTCATATGCGACGCATATGGCTCAGAAATCAACCTCTTTAAATATTT 5158 QY 1362 TACAGAAATTTGGTTTGGTTACCAA 1387 DB 5159 TACAGAGCAGTAATTTGGGAGGCTAA 5184</div></div><div><div>RESULT 12 US-10-087-192-1390 Sequence 1390, Application US/10087192 Publication No. US20020182586A1 GENERAL INFORMATION: APPLICANT: Morris, David W. APPLICANT: Engelhard, Eric K. TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR TITLE OF INVENTION: CANCER FILE REFERENCE: 529452000122 CURRENT APPLICATION NUMBER: US/10/087,192</div></div></div>	<div><div>Publication No. US20020182586A1</div><div>GENERAL INFORMATION: APPLICANT: Engelhard, David W. APPLICANT: Morris, Eric K. TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR TITLE OF INVENTION: CANCER FILE REFERENCE: 529452000122 CURRENT APPLICATION NUMBER: US/10/087,192 CURRENT FILING DATE: 2002-03-01 PRIOR FILING DATE: 2000-12-22 PRIOR FILING DATE: 2001-03-02 NUMBER OF SEQ ID NOS: 2059 SOFTWARE: FastSeq for Windows Version 4.0 SEQ ID NO 1384 LENGTH: 95960 TYPE: DNA ORGANISM: Homo sapiens FEATURE: NAME/KEY: misc feature LOCATION: (1)-(95960) OTHER INFORMATION: n = A,T,C or G US-10-087-192-1384</div></div> <div><div>Query Match 5.3%; Score 201.2; DB 13; Length 95960; Best Local Similarity 72.2%; Pred. No. 1.5e-28; Matches 322; Conservative 0; Mismatches 108; Indels 16; Gaps 4;</div><div><div>QY 954 TGTAGATCACACACTGCTGAGGATTCATCTATGAGACTTTGTCTACATACAGAGAC 1013 DB 4743 TTACTGCTCTATCTCTCTGAGGCTGATGCTGTGAACTTTGTCTTATACAGAAC 4802 QY 1014 CTGGTTTCCAAACC-CCTTTATTTAGCTAAAGCATTTCTTCTACTGACTTTCTTAAG 1072 DB 4803 CTGGTCTCTATAGGCTCCCTCATATTAAGTGAACATTTCTTCTATTGACTTC--AAG 4860 QY 1073 TCTTTAGCAAGCTTAACTCTTTTCAACCAATGCGCAATCAGACAACTTTGAACTACC 1132 DB 4861 TCTTTAGACACAGTTGAACTCTTTTAAACCAATGCGCAATCAGAAATCTTAGACTCCACC 4920 QY 1133 TATGACTGTAGCTCTCTCTGCTTCAAGATCTTTGCTCTTTAAGCTGAACCGATGTGC 1192 DB 4921 TAAGACTGTAA--CCACCCCTCCAAGGTATCTTGTCTTTAGGCTGAATCAATGTAT 4978 QY 1193 ACTTTCCATTTAATGATTTATGCTTTTGTGTTAACTCTCTCCCTTAAATGTATAA 1252 DB 4979 ACCTTCATGTTGACTTATGATTTTACCTAAATTCCTGCTCCCTTAAATGTACAA 5038 QY 1253 AGTAAAGGTGACCTGACCACTCAGGCACACTTTCTCAGCACTCTCTGAGGTGTATCC 1312 DB 5039 ACCAATGTATACCTGACCGCTTTGGGCACACTTTCTCAGGATCTCTTGAGACTGTCCCC 5098 QY 1313 CAGGCCATGGTAAAGT-----CATGTTGGCTCAGAAATCAACCTCTTTAAATATTT 1361 DB 5099 CAGGCCATGGTCACTCATATGCGACGCATATGGCTCAGAAATCAACCTCTTTAAATATTT 5158 QY 1362 TACAGAAATTTGGTTTGGTTACCAA 1387 DB 5159 TACAGAGCAGTAATTTGGGAGGCTAA 5184</div></div><div><div>RESULT 13 US-10-087-192-1372/c Sequence 1372, Application US/10087192 Publication No. US20020182586A1 GENERAL INFORMATION: APPLICANT: Morris, David W. APPLICANT: Engelhard, Eric K. TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR TITLE OF INVENTION: CANCER FILE REFERENCE: 529452000122 CURRENT APPLICATION NUMBER: US/10/087,192 CURRENT FILING DATE: 2002-03-01 PRIOR FILING DATE: 2000-12-22 PRIOR FILING DATE: 2001-03-02 NUMBER OF SEQ ID NOS: 2059 SOFTWARE: FastSeq for Windows Version 4.0 SEQ ID NO 1372</div></div></div>
--	---



```
RESULT 7
US-10-027-632-31701/c
; Sequence 31701, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31701
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31701

Query Match      5.4%; Score 205; DB 13; Length 810;
Best Local Similarity 73.5%; Pred. No. 3.4e-30;
Matches 302; Conservative 0; Mismatches 105; Indels 4; Gaps 3;

Qy 967 CTGCTGAGGATTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTTTCCACA 1026
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1027 ACCCTTTATTTAGCTAAA-GCAATCTTTTCTACTGACTTCTTAAGTCTTTAGCAAAAG 1085
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1086 CTTAACTCTTTCAACCAATGCCAATCAGACAACTTTGAATCTACCTATGACCTGTAAG 1145
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1146 CTCTCTCTGCTTCAAGATCTTGCTCTTTAAGCTGAACCGATGTCACCTTCCATTAA 1205
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1206 TGATTTATGCTTTGCTTGAACCTCTCTCCCTAAAATGTATAAAAGTAAACGGTGAC 1265
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1266 CTGACCACTCAGGCACACTTTCTCAGGACCTCTCAGAGTGTATCCAGGCCATGGTAA 1325
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1326 GTCATGTGGCTCAGATCAACCTCTTAAATATTTTACAGAATTTGGGTT 1376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 421 CTCATATGGCTCTGAATAAACCTCTTTAGATATTTTACAGAGTTTACTTT 371
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-10-027-632-31701/c
; Sequence 31701, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31701
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31701

Query Match      5.4%; Score 205; DB 15; Length 810;
Best Local Similarity 73.5%; Pred. No. 3.4e-30;
Matches 302; Conservative 0; Mismatches 105; Indels 4; Gaps 3;

Qy 967 CTGCTGAGGATTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTTTCCACA 1026
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1027 ACCCTTTATTTAGCTAAA-GCAATCTTTTCTACTGACTTCTTAAGTCTTTAGCAAAAG 1085
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1086 CTTAACTCTTTCAACCAATGCCAATCAGACAACTTTGAATCTACCTATGACCTGTAAG 1145
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1146 CTCTCTCTGCTTCAAGATCTTGCTCTTTAAGCTGAACCGATGTCACCTTCCATTAA 1205
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1206 TGATTTATGCTTTGCTTGAACCTCTCTCCCTAAAATGTATAAAAGTAAACGGTGAC 1265
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1266 CTGACCACTCAGGCACACTTTCTCAGGACCTCTCAGAGTGTATCCAGGCCATGGTAA 1325
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1326 GTCATGTGGCTCAGATCAACCTCTTAAATATTTTACAGAATTTGGGTT 1376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 421 CTCATATGGCTCTGAATAAACCTCTTTAGATATTTTACAGAGTTTACTTT 371
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-10-428-487-11/c
; Sequence 11, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 0980080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
```









Db 61 GGAGAGAAATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAGAAACCGCAGAA 120  
 Qy 121 ATGACTACATTTGAAAGAAATCTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCATCC 180  
 Db 121 ATGACTACATTTGAAAGAAATCTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCATCC 180  
 Qy 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTCGAAGAGTGTCTACACTGTCATT 240  
 Db 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTCGAAGAGTGTCTACACTGTCATT 240  
 Qy 241 AATCACATCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGCTATGAGAACATT 300  
 Db 241 AATCACATCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGCTATGAGAACATT 300  
 Qy 301 GACTCCCTCACAAGAAAGTGAGACAGATTTAGAGAAAGGTGAGAGACAGATATGCCCTT 360  
 Db 301 GACTCCCTCACAAGAAAGTGAGACAGATTTAGAGAAAGGTGAGAGACAGATATGCCCTT 360  
 Qy 361 CTTAGGATCTCTGTTAGTAGGCCCTTGTCTGCGACCCATGAGCATGATTAAGAGTTGTG 420  
 Db 361 CTTAGGATCTCTGTTAGTAGGCCCTTGTCTGCGACCCATGAGCATGATTAAGAGTTGTG 420  
 Qy 421 TTTCCACACTAAATCCTCAAGCTGCTTTATCACCTCCAGCAATGAAGCAATGCAGAA 480  
 Db 421 TTTCCACACTAAATCCTCAAGCTGCTTTATCACCTCCAGCAATGAAGCAATGCAGAA 480  
 Qy 481 TAGCAGACTCTGGCGAAGTTGTTACCCCTGAGCAGTGCATGAACATTTCTTTCTGGCTA 540  
 Db 481 TAGCAGACTCTGGCGAAGTTGTTACCCCTGAGCAGTGCATGAACATTTCTTTCTGGCTA 540  
 Qy 541 AAGTTTAGAATATATCTTATATATATATATATCTTAGGCAACTCTGATATGGGCATCTG 600  
 Db 541 AAGTTTAGAATATATCTTATATATATATATCTTAGGCAACTCTGATATGGGCATCTG 600  
 Qy 601 TGGCTTAGTGAAATCATAGAAATGACAAATGACCTTAAATATTTCTATGTTTTTGC 660  
 Db 601 TGGCTTAGTGAAATCATAGAAATGACAAATGACCTTAAATATTTCTATGTTTTTGC 660  
 Qy 661 TTGTAAGTTTGAGGACATGGAGTGATATAAAAAAACTTTCTTAGGCAATTAATGTAAA 720  
 Db 661 TTGTAAGTTTGAGGACATGGAGTGATATAAAAAAACTTTCTTAGGCAATTAATGTAAA 720  
 Qy 721 TGAATAATAATTTCTAATCCCCCTGACTAATGATGAGACCTCTTAGGCCAAAGAGA 780  
 Db 721 TGAATAATAATTTCTAATCCCCCTGACTAATGATGAGACCTCTTAGGCCAAAGAGA 780  
 Qy 781 CCTCAGATGAACCTGAAAGACTGAAATCTGGCCATGATAGGAAGGAGTGAGACACACC 840  
 Db 781 CCTCAGATGAACCTGAAAGACTGAAATCTGGCCATGATAGGAAGGAGTGAGACACACC 840  
 Qy 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAACAAGTGACAGGATGAGTCAAGACT 900  
 Db 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAACAAGTGACAGGATGAGTCAAGACT 900  
 Qy 901 GATCAATATAGACTGATGTCGCAATAGAGTCCCAATTCCAACCTGACTCTGGGTAGAT 960  
 Db 901 GATCAATATAGACTGATGTCGCAATAGAGTCCCAATTCCAACCTGACTCTGGGTAGAT 960  
 Qy 961 CACACACTGTCTGAGGAGTTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGTT 1020  
 Db 961 CACACACTGTCTGAGGAGTTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGTT 1020  
 Qy 1021 TCCCAACCCCTTTATTTTAGCTAAGCAATTTCTTTCTACTGACTCTTTAAAGTCTTTAGA 1080  
 Db 1021 TCCCAACCCCTTTATTTTAGCTAAGCAATTTCTTTCTACTGACTCTTTAAAGTCTTTAGA 1080  
 Qy 1081 CAAGCTTAACTCTTCAACCAATTTGCCAATCAGACAAACTTTGAATCTACCTAGACCT 1140  
 Db 1081 CAAGCTTAACTCTTCAACCAATTTGCCAATCAGACAAACTTTGAATCTACCTAGACCT 1140  
 Qy 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAACCGATGTCATTTCCA 1200

Db 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAACCGATGTCATTTCCA 1200  
 Qy 1201 TTTAATGATTTATGTTCTTCTTGTAACTCCTGTCTCCCTTAAATATGATAAAAGTAAACG 1260  
 Db 1201 TTTAATGATTTATGTTCTTCTTGTAACTCCTGTCTCCCTTAAATATGATAAAAGTAAACG 1260  
 Qy 1261 GTGACCTGACACCTCTAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCAGGCCAT 1320  
 Db 1261 GTGACCTGACACCTCTAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320  
 Qy 1321 GGTAAAGTCATGTTGGCTCAGAACTCACTCTTTTAAATATTTTACAGAAATTTGGTTTGG 1380  
 Db 1321 GGTAAAGTCATGTTGGCTCAGAACTCACTCTTTTAAATATTTTACAGAAATTTGGTTTGG 1380  
 Qy 1381 TTACCAATAAGTCTCCACAAATATATGTCACAAATCTTCAATTTCCAAGCTCTGCTACCA 1440  
 Db 1381 TTACCAATAAGTCTCCACAAATATATGTCACAAATCTTCAATTTCCAAGCTCTGCTACCA 1440  
 Qy 1441 AATTTCAATGCCAATCTCCCGATCCCAATTAACCTATTTTCACTTTTGAAGTGTAAATCTA 1500  
 Db 1441 AATTTCAATGCCAATCTCCCGATCCCAATTAACCTATTTTCACTTTTGAAGTGTAAATCTA 1500  
 Qy 1501 CTCAATAAACTGTGTAAAGCAGTGCAGACCTTTGCTAACTGACATTTACTTCAAT 1560  
 Db 1501 CTCAATAAACTGTGTAAAGCAGTGCAGACCTTTGCTAACTGACATTTACTTCAAT 1560  
 Qy 1561 TTTTCTTTTCTATGTACTGGATATTTTTCATATATAAACTTGCAATTAAGTTCAAAAAAT 1620  
 Db 1561 TTTTCTTTTCTATGTACTGGATATTTTTCATATATAAACTTGCAATTAAGTTCAAAAAAT 1620  
 Qy 1621 TAATAGTTTTTGACATTTGCTTCTGAGAGAGAAATGAAAGTGTGCAAAAAATAAAA 1680  
 Db 1621 TAATAGTTTTTGACATTTGCTTCTGAGAGAGAAATGAAAGTGTGCAAAAAATAAAA 1680  
 Qy 1681 AAGATGAATGAACATATATATTTTCAATTTTTTCAATTTTCTAGTCAACAGAGAAATC 1740  
 Db 1681 AAGATGAATGAACATATATATTTTCAATTTTTTCAATTTTCTAGTCAACAGAGAAATC 1740  
 Qy 1741 GAAGGATTTCTGTTCAATAATAGTAAAAATGAAAAATAAACTTGTGCTTATATTTCTTT 1800  
 Db 1741 GAAGGATTTCTGTTCAATAATAGTAAAAATGAAAAATAAACTTGTGCTTATATTTCTTT 1800  
 Qy 1801 GCAACACTAGTTAAATTTAACTGCTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860  
 Db 1801 GCAACACTAGTTAAATTTAACTGCTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860  
 Qy 1861 TTCTGTTTTTAAAAATCAAGCAAACTGAAAAATAATCCATCTAATTAATGCTTTCTTTCCC 1920  
 Db 1861 TTCTGTTTTTAAAAATCAAGCAAACTGAAAAATAATCCATCTAATTAATGCTTTCTTTCCC 1920  
 Qy 1921 AAGAGTTTTTTAATGATGCGAGCTTCTAAATTTTGGAGACAAAAAGCCTTAATGACAA 1980  
 Db 1921 AAGAGTTTTTTAATGATGCGAGCTTCTAAATTTTGGAGACAAAAAGCCTTAATGACAA 1980  
 Qy 1981 TGCAATTCATTTATATTTTTTTGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
 Db 1981 TGCAATTCATTTATATTTTTTTGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
 Qy 2041 AGATGCTTGGGACAGAGAGTTTGGATTTTCAAGTTTATTTTGGATTTTGGAAATATTT 2100  
 Db 2041 AGATGCTTGGGACAGAGAGTTTGGATTTTCAAGTTTATTTTGGATTTTGGAAATATTT 2100  
 Qy 2101 CCATACATATAATAGAGAGTTTGGAAAAATGGGATTCAGTCTAATCATATAAAATTCACCTTA 2160  
 Db 2101 CCATACATATAATAGAGAGTTTGGAAAAATGGGATTCAGTCTAATCATATAAAATTCACCTTA 2160  
 Qy 2161 TGTGATATACACCTTATCTGATAGCCTGAGGTAATTTTATACATATTTTAAATAA 2220  
 Db 2161 TGTGATATACACCTTATCTGATAGCCTGAGGTAATTTTATACATATTTTAAATAA 2220  
 Qy 2221 TTTTATCCCTGAAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAGAGTGTCACTAT 2280  
 Db 2221 TTTTATCCCTGAAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAGAGTGTCACTAT 2280

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 18:25:09 ; Search time 1741 Seconds  
(without alignments)  
11676.296 Million cell updates/sec

Title: US-10-005-907-1

Perfect score: 3762

Sequence: 1 ggaacacgagtcactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 270181610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3762	100.0	3762	15	US-10-005-907-1
C 2	234.4	6.2	191584	17	Sequence 1, Appli
C 3	224.6	6.0	830	13	Sequence 800, App
C 4	224.6	6.0	830	13	Sequence 125465,
C 5	207.4	5.5	58320	17	Sequence 125465,
C 6	207.4	5.5	63313	13	Sequence 88, Appl
C 7	205	5.4	810	13	Sequence 526, App
C 8	205	5.4	810	15	Sequence 31701, A
C 9	203.8	5.4	301692	16	Sequence 31701, A
C 10	203.8	5.4	310268	17	Sequence 11, Appl
C 11	201.2	5.3	95960	13	Sequence 195, App
C 12	201.2	5.3	95960	13	Sequence 1384, App
					Sequence 1390, App

C 13	201.2	5.3	130207	13	US-10-087-192-1372	Sequence 1372, Ap
C 14	200.4	5.3	17239	9	US-09-764-869-2326	Sequence 2326, Ap
C 15	200.4	5.3	17239	14	US-10-091-504-2326	Sequence 2326, Ap
C 16	200.4	5.3	17239	16	US-10-027-577-2326	Sequence 2, Appli
C 17	199.8	5.3	181259	16	US-10-456-930-2	Sequence 274, App
C 18	199.6	5.3	73145	13	US-10-087-192-274	Sequence 19728, A
C 19	199.2	5.3	728	13	US-10-027-632-19728	Sequence 19728, A
C 20	199.2	5.3	728	15	US-10-027-632-19728	Sequence 2, Appli
C 21	197.4	5.2	290	10	US-09-854-867-2	Sequence 6056, App
C 22	197.4	5.2	1743	10	US-09-764-891-6056	Sequence 236, App
C 23	197.4	5.2	1743	15	US-10-091-414-236	Sequence 6057, App
C 24	197.4	5.2	1746	10	US-09-764-891-6057	Sequence 237, App
C 25	197.4	5.2	1746	15	US-10-091-414-237	Sequence 86, Appli
C 26	197.4	5.2	6686	9	US-09-736-960-86	Sequence 46, Appli
C 27	197	5.2	166043	16	US-10-235-192A-46	Sequence 8396, App
C 28	196.8	5.2	27154	10	US-09-764-891-8396	Sequence 169, App
C 29	196.2	5.2	11750	17	US-10-367-094-189	Sequence 105966,
C 30	196	5.2	586	13	US-10-027-632-105966	Sequence 105966,
C 31	196	5.2	586	15	US-10-027-632-105966	Sequence 1, Appli
C 32	195.4	5.2	288	14	US-10-115-278-1	Sequence 1486, App
C 33	194.8	5.2	54169	13	US-10-087-192-1486	Sequence 2670, App
C 34	193.8	5.2	8207	9	US-09-764-877-2670	Sequence 741, App
C 35	193.8	5.2	8207	10	US-09-764-872-741	Sequence 2670, App
C 36	193.8	5.1	598	13	US-10-242-515-2670	Sequence 269156,
C 37	193.6	5.1	598	15	US-10-027-632-269156	Sequence 269156,
C 38	193.6	5.1	598	15	US-10-027-632-269156	Sequence 29373, A
C 39	193.6	5.1	628	13	US-10-027-632-29373	Sequence 29373, A
C 40	193.6	5.1	628	15	US-10-027-632-29373	Sequence 265848,
C 41	193.2	5.1	1599	13	US-10-027-632-265848	Sequence 265848,
C 42	193.2	5.1	1599	15	US-10-027-632-265848	Sequence 76, Appli
C 43	193.2	5.1	96588	15	US-10-085-117-76	Sequence 11, Appli
C 44	193.2	5.1	185695	14	US-10-020-141-11	Sequence 1, Appli
C 45	193.2	5.1	185695	14	US-10-017-721-1	

#### ALIGNMENTS

RESULT 1

US-10-005-907-1

; Sequence 1, Application US/10005907

; Publication No. US20030166881A1

; GENERAL INFORMATION:

; APPLICANT: Union Chimique Belge, S.A.

; APPLICANT: No. US20030166881A1, Karl

; APPLICANT: Pirozzi, Gregory

; APPLICANT: Einstein, Richard

; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST

; FILE REFERENCE: 053529-5005

; CURRENT APPLICATION NUMBER: US/10/005,907

; CURRENT FILING DATE: 2001-12-07

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: Patent version 3.1

; SEQ ID NO 1

; LENGTH: 3762

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (25)..(432)

; OTHER INFORMATION:

US-10-005-907-1

Query Match 100.0%; Score 3762; DB 15; Length 3762;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACCTGTGAAAAGATGGGAAATATCTCTCGGAAAACCTCAGTTGCTG 60

Db 1 GAGAAACCGAGTCACCTGTGAAAAGATGGGAAATATCTCTCGGAAAACCTCAGTTGCTG 60

QY 61 GGAGAGATCAAAAGAACCCCAAGAAAGAGATGAGGAAAGAAACGGCAGGAA 120

Page Blank (uspto)

Query Match 10.4%; Score 392.2; DB 5; Length 437;  
 Best Local Similarity 99.2%; Pred. No. 1.5e-47;  
 Matches 394; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1571	CTATGCTACTGGATATTTTTCGATATATAAACTTGCAGTAATAGTTCAAAAAATTAATAGTTTTT	1630
Db	41	CGAGGTACTGGATATTTTTCGATATATAAACTTGCAGTAATAGTTCAAAAAATTAATAGTTTTT	100
QY	1631	TGACATTTGGCTTTTCTGAGAGAGAAATTCAAAAGTGTCAAAAATTAATAAAGATGAAAT	1690
Db	101	TGACATTTGGCTTTTCTGAGAGAGAAATTCAAAAGTGTCAAAAATTAATAAAGATGAAAT	160
QY	1691	GAAGCATATATAATTTGTCAATTTTTCAAATTTTCTAGTCAACAGAGAAATCGAAGGATTC	1750
Db	161	GAAGCATATATAATTTGTCAATTTTTCAAATTTTCTAGTCAACAGAGAAATCGAAGGATTC	220
QY	1751	GTTCAAAATATTAGTAAAAATTTGAAAATAAATCTTGTGCTTATATTTTGTTCGAAACACT	1810
Db	221	GTTCAAAATATTAGTAAAAATTTGAAAATAAATCTTGTGCTTATATTTTGTTCGAAACACT	280
QY	1811	AGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGTGTAGTTTCTGGTTTTT	1870
Db	281	AGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGTGTAGTTTCTGGTTTTT	340
QY	1871	AAAAATTCAGCAAACTGGAAAAATTAATCCATCTAATTAATGCTTTTCTCCCAAGAGTTTT	1930
Db	341	AAAAATTCAGCAAACTGGAAAAATTAATCCATCTAATTAATGCTTTTCTCCCAAGAGTTTT	400
QY	1931	TTAATGATATGCCAGCTTCCTAATTTTGGAGACAAAAG	1967
Db	401	TTAATGATATGCCAGCTTCCTAATTTTGGAGACAAAAG	437

Search completed: November 18, 2004, 14:08:43  
 Job time : 1858 secs



```
Db      376 TTAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATTG 421
RESULT 14
ABV31609
ID      ABV31609 standard; cDNA; 437 BP.
XX      AC      ABV31609;
XX      DT      16-SEP-2002 (first entry)
XX      DE      Human prostate expression marker cDNA 31600.
XX      KW      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX      KW      pharmacogenomic marker; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200160860-A2.
XX      PD      23-AUG-2001.
XX      PF      20-FEB-2001; 2001WO-US005171.
XX      PR      17-FEB-2000; 2000US-0183319P.
XX      PR      16-MAR-2000; 2000US-0189862P.
XX      PR      25-MAY-2000; 2000US-0207454P.
XX      PR      09-JUN-2000; 2000US-0211314P.
XX      PR      18-JUL-2000; 2000US-0219007P.
XX      PR      13-DEC-2000; 2000US-0255281P.
XX      PA      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX      PI      Schlegel R, Endege WO, Monahan JE;
XX      DR      MPI; 2001-662795/76.
XX      PT      Novel isolated nucleic acid molecule associated with cancerous state of
XX      PT      prostate cells and correlating with presence of prostate cancer, useful
XX      PS      Claim 1; Page 6790; 11750pp; English.
XX      CC      The invention relates to an isolated nucleic acid molecule (I) comprising
XX      CC      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX      CC      specification or its complement. (I) is useful for: (a) assessing whether
XX      CC      a patient is afflicted with prostate cancer; (b) monitoring the
XX      CC      progression of prostate cancer in a patient; (c) assessing the efficacy
XX      CC      of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX      CC      the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX      CC      (e) selecting a composition for inhibiting prostate cancer in a patient;
XX      CC      (f) assessing the prostate cell carcinogenic potential of a compound;
XX      CC      (g) determining whether prostate cancer has metastasized in a patient;
XX      CC      (h) assessing the aggressiveness or indolence of prostate cancer in a patient
XX      CC      ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX      SQ      Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;

Query Match      10.4%; Score 392.2; DB 5; Length 437;
Best Local Similarity 99.2%; Pred. No. 1.5e-47;
Matches 394; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1571 CTATGCTAGTATTTTGGCATATAAACTTGCAGTAAATAGTTCAAAATTAATAGTTTT 1630
Db      41 CGAGGTACTGGATATTTTTCATATAAACTTGCAGTAAATAGTTCAAAATTAATAGTTTT 100
QY      1631 TGACATTCGCTTTTCTGAGAGAGAAATTTGAAGTGCACAAATATAAAAGATGAAAT 1690
Db      101 TGACATTCGCTTTTCTGAGAGAGAAATTTGAAGTGCACAAATATAAAAGATGAAAT 160
QY      1691 GAAGCATATATAATGTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGATTCT 1750
Db      161 GAAGCATATATAATGTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGATTCT 220

1751 GTTCAAAATATTAGTAAAAATTGAAATAAACTTGTGCTTAATATTTGTTTGGCAACACACT 1810
221 GTTCAAAATATTAGTAAAAATTGAAATAAACTTGTGCTTAATATTTGTTTGGCAACACACT 280
1811 AGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGGTGGATGTGTAGTTTCTGGTTTT 1870
281 AGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGGTGGATGTGTAGTTTCTGGTTTT 340
1871 AAAATTCAAGCAAACTGGAAAAATAATCCATCTAATTAATGCTTTTCCCAAGAAGTTTT 1930
341 AAAATTCAAGCAAACTGGAAAAATAATCCATCTAATTAATGCTTTTCCCAAGAAGTTTT 400
1931 TTAATGATATGCCAGCTTCTTAATTTGGAGACAAAAG 1967
401 TTAATGATATGCCAGCTTCTTAATTTGGAGACAAAAG 437

RESULT 15
ABV40245
ID      ABV40245 standard; cDNA; 437 BP.
XX      AC      ABV40245;
XX      DT      16-SEP-2002 (first entry)
XX      DE      Human prostate expression marker cDNA 40236.
XX      KW      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX      KW      pharmacogenomic marker; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200160860-A2.
XX      PD      23-AUG-2001.
XX      PF      20-FEB-2001; 2001WO-US005171.
XX      PR      17-FEB-2000; 2000US-0183319P.
XX      PR      16-MAR-2000; 2000US-0189862P.
XX      PR      25-MAY-2000; 2000US-0207454P.
XX      PR      09-JUN-2000; 2000US-0211314P.
XX      PR      18-JUL-2000; 2000US-0219007P.
XX      PR      13-DEC-2000; 2000US-0255281P.
XX      PA      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX      PI      Schlegel R, Endege WO, Monahan JE;
XX      DR      MPI; 2001-662795/76.
XX      PT      Novel isolated nucleic acid molecule associated with cancerous state of
XX      PT      prostate cells and correlating with presence of prostate cancer, useful
XX      PS      Claim 1; Page 8129; 11750pp; English.
XX      CC      The invention relates to an isolated nucleic acid molecule (I) comprising
XX      CC      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX      CC      specification or its complement. (I) is useful for: (a) assessing whether
XX      CC      a patient is afflicted with prostate cancer; (b) monitoring the
XX      CC      progression of prostate cancer in a patient; (c) assessing the efficacy
XX      CC      of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX      CC      the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX      CC      (e) selecting a composition for inhibiting prostate cancer in a patient;
XX      CC      (f) assessing the prostate cell carcinogenic potential of a compound;
XX      CC      (g) determining whether prostate cancer has metastasized in a patient;
XX      CC      (h) assessing the aggressiveness or indolence of prostate cancer in a patient
XX      CC      ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX      SQ      Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;
```

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX PA Schlegel R, Endege WO, Monahan JE;  
 XX PI WPI; 2001-662795/76.  
 XX DR  
 XX  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 XX Claim 1; Page 1680; 11750pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 XX Sequence 397 BP; 142 A; 53 C; 64 G; 138 T; 0 U; 0 Other;  
 XX  
 Query Match 10.5%; Score 393.8; DB 5; Length 397;  
 Best Local Similarity 99.5%; Pred. No. 8.8e-48;  
 Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1573 ATGCTACTGATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAAATAGTTTGG 1632  
 Db 1 AGTACTGGATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAAATAGTTTGG 60  
 QY 1633 ACATTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAATGA 1692  
 Db 61 ACATTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAATGA 120  
 QY 1693 AGCATATATAATGTCAAATTTTCAATTTTCTAGTCAACAGAGATCGAAGATTCGTGT 1752  
 Db 121 AGCATATATAATGTCAAATTTTCAATTTTCTAGTCAACAGAGATCGAAGATTCGTGT 180  
 QY 1753 TCAATATATTAGTAAATAATGAAATAAACTTGTGCTTATATTTTTCACACACTAG 1812  
 Db 181 TCAATATATTAGTAAATAATGAAATAAACTTGTGCTTATATTTTTCACACACTAG 240  
 QY 1813 TTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTTAA 1872  
 Db 241 TTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTTAA 300  
 QY 1873 AATTCAGCAACTGGAATAATTCATCTAATATGCTTTCTTCCCAAGAGATTTT 1932  
 Db 301 AATTCAGCAACTGGAATAATTCATCTAATATGCTTTCTTCCCAAGAGATTTT 360  
 QY 1933 AATGATATCCAGCTTCCCTAAATTTGGAGCAAAAAGCC 1969  
 Db 361 AATGATATCCAGCTTCCCTAAATTTGGAGCAAAAAGCC 397  
 RESULT 13  
 ABV31276  
 ID ABV31276 standard; cDNA; 421 BP.  
 XX AC ABV31276;  
 XX  
 XX 16-SEP-2002 (first entry)  
 DT  
 DE Human prostate expression marker cDNA 31267.  
 XX  
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.  
 OS  
 XX WO200160860-A2.  
 PN  
 XX 23-AUG-2001.  
 PD  
 XX  
 XX 20-FEB-2001; 2001WO-US005171.  
 PF  
 XX 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA  
 XX Schlegel R, Endege WO, Monahan JE;  
 PI  
 XX WPI; 2001-662795/76.  
 DR  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 XX Claim 1; Page 6740; 11750pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 XX Sequence 421 BP; 143 A; 61 C; 71 G; 145 T; 0 U; 1 Other;  
 XX  
 Query Match 10.4%; Score 392.2; DB 5; Length 421;  
 Best Local Similarity 97.8%; Pred. No. 1.5e-47;  
 Matches 397; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 1571 CTATGCTACTGATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAAATAGTTT 1630  
 Db 16 CGAGGTACTTTATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAAATAGTTT 75  
 QY 1631 TGACATTGGCTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAAT 1690  
 Db 76 TGACATTGGCTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAAT 135  
 QY 1691 GAAGCATATATAATTTGTCAAATTTTCTAGTCAACAGAGATCGAAGATTCCT 1750  
 Db 136 GAAGCATATATAATTTGTCAAATTTTCTAGTCAACAGAGATCGAAGATTCCT 195  
 QY 1751 GTTCAAAATATTAGTAAAAATTGAAAATAAACTTGTGCTTATATTTTGTGCAACACACT 1810  
 Db 196 GTTCAAAATATTAGTAAAAATTGAAAATAAACTTGTGCTTATATTTTGTGCAACACACT 255  
 QY 1811 AGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTT 1870  
 Db 256 AGTCAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTT 315  
 QY 1871 AAAATTTCAAGCAAACTGGAAAAATAATCCATCTAATTAATGCTTTCTTCCCAAGAGATTT 1930  
 Db 316 AAAATTTCAAGCAAACTGGAAAAATAATCCATCTAATTAATGCTTTCTTCCCAAGAGATTT 375  
 QY 1931 TTAATGATATGCCAGCTTCCCTAAATTTGGAGCAAAAAGCCCTTAATTG 1976

Db 222 GGAGAGAATCAAAGAACCCCAAGAAAGGAACCCAGATGAGGAAGAAAGGACGAGAA 281  
 Qy 121 ATGACTAGATTGAAGAAAGAACTTCAAGATCAAGATAAGAAAGCAAGAAAGTTTCATCC 180  
 Db 282 ATGACTAGATTGAAGAAAGAACTTCAAGATCAAGATAAGAAAGCAAGAAAGTTTCATCC 341  
 Qy 181 ACTTCTAATCAGGAAACGAGAATGGCAGTGGTCTGGAAGAGTGTGCTACACTGTCAAT 240  
 Db 342 ACTTCTAATCAGGAAACGAGAATGGCAGTGGTCTGGAAGAGTGTGCTACACTGTCAAT 401  
 Qy 241 AATCACAATCCCAATCAGAGATCCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACATT 300  
 Db 402 AATCACAATCCCAATCAGAGATCCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACATT 461  
 Qy 301 GACTCCCTCACAAGAAAGTGAAGAGTGTGAGAGAAAGTGAAGAGTGTGAGAGTGTGAGAGTGTG 360  
 Db 462 GACTCCCTCACAAGAAAGTGAAGAGTGTGAGAGAAAGTGAAGAGTGTGAGAGTGTGAGAGTGTG 521  
 Qy 361 CTTAGGACTTCTGTTAGTAGGCTTGTTCCTGACCCATGAGCATGATTATGAAGTTGTG 420  
 Db 522 CTTAGGACTTCTGTTAGTAGGCTTGTTCCTGACCCATGAGCATGATTATGAAGTTGTG 581  
 Qy 421 TTTCCACACTAAATCTCAAGCTGCTTTATCACCCTCCAGCAATGAAGACAATGCAGAA 480  
 Db 582 TTTCCACACTAAATCTCAAGCTGCTTTATCACCCTCCAGCAATGAAGACAATGCAGAA 641  
 Qy 481 TAGCAGACTCTGGCGAAGTTGTTAC 506  
 Db 642 TAGCAGACTCTGGCGAAGTTGTTAC 667

RESULT 11

ABV01270

ID ABV01270 standard; cDNA; 433 BP.

XX AC ABV01270;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 1261.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 295; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX

SQ Sequence 433 BP; 146 A; 64 C; 79 G; 144 T; 0 U; 0 Other;

Query Match 10.5%; Score 394.2; DB 5; Length 433;

Best Local Similarity 99.2%; Pred. No. 7.6e-48;

Matches 396; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1571 CTATGTAATCGATATTTTGCATATAAATTCGAGTAAATAGTTCAAAAATTAATAGTTT 1630

Db 35 CGAGTACTGGATATTTTGCATATAAATTCGAGTAAATAGTTCAAAAATTAATAGTTT 94

Qy 1631 TGACATTTGGCTTTTCTGAGAGAGAGAAATTCGAGTAAATAGTTCAAAAATTAATAGTTT 1690

Db 95 TGACATTTGGCTTTTCTGAGAGAGAGAAATTCGAGTAAATAGTTCAAAAATTAATAGTTT 154

Qy 1691 GAAGCATATATAATGTCAATTTTCAATTTTCTAGTCAACAGAGAAATCGAAGATTCT 1750

Db 155 GAAGCATATATAATGTCAATTTTCAATTTTCTAGTCAACAGAGAAATCGAAGATTCT 214

Qy 1751 GTTCAATATATAGTAAATTTGAAATTAATAGTTCGCTTATATTTTTCGACACACT 1810

Db 215 GTTCAATATATAGTAAATTTGAAATTAATAGTTCGCTTATATTTTTCGACACACT 274

Qy 1811 AGTTAATTTAACCTGTGACTAGTTATCTCTACCGAAGTGGATGTGTAGTTTCTGGTTT 1870

Db 275 AGTTAATTTAACCTGTGACTAGTTATCTCTACCGAAGTGGATGTGTAGTTTCTGGTTT 334

Qy 1871 AAAATTCAGCAAACTCGAAAAATTAATCCATCTAATATGCTTTTCCCAAGAGTTT 1930

Db 335 AAAATTCAGCAAACTCGAAAAATTAATCCATCTAATATGCTTTTCCCAAGAGTTT 394

Qy 1931 TTAATGATATGCCAGCTTCCCTAATTTGGAGACAAGGCC 1969

Db 395 TTAATGATATGCCAGCTTCCCTAATTTGGAGACAAGGCC 433

RESULT 12

ABV10439

ID ABV10439 standard; cDNA; 397 BP.

XX AC ABV10439;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 10430.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

Db 303 ATCGAAGGATCTGTTCAAAATATTAGTAAATAGAAATGAAATCTGCTATATTTG 362  
Qy 1798 TTTGCAACACACTAGTATTAATTTAACTGTGACTAGTATCTCTACCGAAGTGGATGTG 1857  
Db 363 TTTGCAACACACTAGTATTAATTTAACTGTGACTAGTATCTCTACCGAAGTGGATGTG 422  
Qy 1858 AGTTTCTGTTTAAATTTCAAGCAAACTGGAAATTAATCAATCTAATTTATGCTTCTTT 1917  
Db 423 AGTTTCTGTTTAAATTTCAAGCAAACTGGAAATTAATCAATCTAATTTATGCTTCTTT 482  
Qy 1918 CCCAAGAGTGTTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATCA 1977  
Db 483 CCCAAGAGTGTTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATCA 542  
Qy 1978 CAATGATTCATTATATATTTTTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 2037  
Db 543 CAATGATTCATTATATATTTTTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 602  
Qy 2038 ATGAGATGCTTGGACCAAGTGTTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTA 2097  
Db 603 ATGAGATGCTTGGACCAAGTGTTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTA 662  
Qy 2098 TTTCCATACATATAATGAGAGAGTTTGGAAAATGGGATTCAGTCTAATCATAAAATTCAC 2157  
Db 663 TTTCCATACATATAATGAGAGAGTTTGGAAAATGGGATTCAGTCTAATCATATAGATGCAC 722  
Qy 2158 TTATGTTTGGATATACACCTTATCTGAATAGCCTGAAAGTAAATTTTATACATATTTTAAA 2217  
Db 723 TTATGTTTGGATATACACCTTATCTGAATAGCCTGAAAGTAAATTTTATACATATTTTAAA 782  
Qy 2218 TAAATTTATGCTGAAACAGAGTTTGGCACATTTGACCATCATGAAAGCAGAGTGTCTAC 2277  
Db 783 TAAATTTATGCTGAAACAGAGTTTGGCACATTTGACCATCATGAAAGCAGAGTGTCTAC 842  
Qy 2278 TATTTCAAGTCAGTGTCTAAAAGTTTTCAGATGTTAAGCTGGTGTGATGTCAGTTTCATGCCAG 2337  
Db 843 TATTTCAAGTCAGTGTCTAAAAGTTTTCAGATGTTAAGCTGGTGTGATGTCAGTTTCATGCCAG 902  
Qy 2338 TGATCCGAGTACTTTGGGAAGCCAGACAGAGTGGATCTCTTTGAGCCAGGAGTTTGAAGC 2397  
Db 903 TGATCCGAGTACTTTGGGAAGCCAGACAGAGTGGATCTCTTTGAGCCAGGAGTTTGAAGC 962  
Qy 2398 CAGACTGCAACACAGTGAAGCTGTTTCTCAATATTAATTAATAATTTAGCCAGGTGT 2457  
Db 963 CAGACTGCAACACAGTGAAGCTGTTTCTCAATATTAATTAATAATTTAGCCAGGTGT 1022  
Qy 2458 GGTGGTGCACACCTGTAGTCCAGGCTACTCAGGAGGCTGAGG-TAGTAGGATTTGTTGAG 2516  
Db 1023 GGTGGTGCACACCTGTAGTCCAGGCTACTCAGGAGGCTGAGG-TAGTAGGATTTGTTGAG 1082  
Qy 2517 ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 2560  
Db 1083 ACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 1126

RESULT 10  
ID ADC30760  
XX ADC30760 standard; cdna; 667 BP.

AC ADC30760;

XX 18-DEC-2003 (first entry)

XX Human novel cdna sequence, SEQ ID NO:842.

XX Human; diagnostic; drug screening; forensics; gene mapping;  
XX biodiversity assessment; parkinson's disease; Alzheimer's disease;  
XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
XX ulcers; osteoporosis; autoimmune disease; cancer;  
XX molecular weight marker; food supplement; antiparkinsonian; nootropic;  
XX neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;  
XX antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

gene therapy; chromosome 1; gene; ss.  
XX Homo sapiens.  
XX WO2003029271-A2.  
XX 10-APR-2003.  
XX 24-SEP-2002; 2002WO-US030474.  
XX 24-SEP-2001; 2001US-0324631P.  
XX (HYSE-) HYSEQ INC.  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
XX Haley-Vicente D, Drmanac RT;  
XX WPI; 2003-371981/35.  
XX P-PSDB; ADC31731.  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
XX treating conditions such as neurodegenerative diseases, anemias, platelet  
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
XX cancer.  
XX Claim 1; SEQ ID NO 842; 1185pp; English.  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
XX ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The  
XX invention also relates to nucleic acid sequences over 99% identical with  
XX the novel human cDNAs. The invention additionally encompasses expression  
XX vectors and host cells comprising a nucleic acid of the invention; the  
XX recombinant production of a polypeptide of the invention; an antibody  
XX against a polypeptide of the invention; a method of detecting  
XX polynucleotides or polypeptides of the invention; and methods of  
XX identifying a compound which binds to a polypeptide of the invention. The  
XX invention further discloses methods of preventing, treating or  
XX ameliorating a medical condition; kits comprising polynucleotide probes  
XX and/or monoclonal antibodies for carrying out the methods of the  
XX invention; methods for the identification of compounds that modulate the  
XX expression or activity of the polynucleotide and/or polypeptide; and 767  
XX contig sequences corresponding to the cDNA sequences of the invention  
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
XX -ADC33394). The nucleic acids and polypeptides of the invention are  
XX useful in diagnostics, drug screening, forensics, gene mapping, in the  
XX identification of mutations responsible for genetic disorders or other  
XX traits, for assessing biodiversity, and in producing many other types of  
XX data and products dependent on DNA and amino acid sequences. They are  
XX also used for treating diseases such as Parkinson's disease, Alzheimer's  
XX disease and other neurodegenerative diseases, anaemia, platelet  
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
XX cancer. The nucleic acids may also be used as hybridisation probes or  
XX primers, and in the recombinant production of a protein. The polypeptides  
XX are also useful in generating antibodies, as molecular weight markers,  
XX and as food supplements. The present sequence represents a specifically  
XX claimed human cDNA sequence of the invention. Note: The sequence data for  
XX this patent did not form part of the printed specification, but was  
XX obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Query Match 13.5%; Score 506; DB 10; Length 667;

Best Local Similarity 100.0%; Pred. No. 6.8e-64;

Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAAACCGAGTCACCTGTGAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCTG 60

Db 162 GAGAACCGAGTCACCTGTGAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCTG 221

Qy 61 GGAGAGATCAAAAGAACGCCAAGAGGAACCCAGATGAGGAAGAAACGCGACGAA 120

QY 1618 AATTAATAGTTTTTGGACATGGCTTTTCTGAGAGAGAAATGAAAGTGTCAAAAATAA 1677  
 Db 183 AATTAATAGTTTTTGGACATGGCTTTTCTGAGAGAGAAATGAAAGTGTCAAAAATAA 242  
 QY 1678 AAAAGATGAATGAAGCATATATAATGTCAATTTTTTCAATTTCTAGTCAACAGAGA 1737  
 Db 243 AAAAGATGAATGAAGCATATATAATGTCAATTTTTTCAATTTCTAGTCAACAGAGA 302  
 QY 1738 ATCGAAGATTCTGTTCAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATG 1797  
 Db 303 ATCGAAGATTCTGTTCAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATG 362  
 QY 1798 TTTCGACACACTAGTTAAATTTACCTGTGCTAGTATCTCTACCGAGAGTGTGTGT 1857  
 Db 363 TTTCGACACACTAGTTAAATTTACCTGTGCTAGTATCTCTACCGAGAGTGTGTGT 422  
 QY 1858 AGTTCTGGTTTTAAATTTCAAGCAAACTGGAAATATATCCATCTAAATATGCTTTCTTT 1917  
 Db 423 AGTTCTGGTTTTAAATTTCAAGCAAACTGGAAATATATCCATCTAAATATGCTTTCTTT 482  
 QY 1918 CCCAAGAGTTTTTAAATGATGACGCTTCTTAATTTGGAGACAAAAGCCTTAATGA 1977  
 Db 483 CCCAAGAGTTTTTAAATGATGACGCTTCTTAATTTGGAGACAAAAGCCTTAATGA 542  
 QY 1978 CAATGCATCTATTATATATTTTTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 2037  
 Db 543 CAATGCATCTATTATATATTTTTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 602  
 QY 2038 ATGAGATGCTTGGACAGAGTGTTTTGGATTTTCAAGTTTATTTTGGATTTTGGATA 2097  
 Db 603 ATGAGATGCTTGGACAGAGTGTTTTGGATTTTCAAGTTTATTTTGGATTTTGGATA 662  
 QY 2098 TTTCATACATATAATGAGAGAGTTGGAATGGATTCAGTCTAATCAATPAAATTCAC 2157  
 Db 663 TTTCATACATATAATGAGAGAGTTGGAATGGATTCAGTCTAATCAATPAAATTCAC 722  
 QY 2158 TTATGTTTGATATACACTTCTGAATAGCTCAAGTATATTTTATACATATTTTAAA 2217  
 Db 723 TTATGTTTGATATACACTTCTGAATAGCTCAAGTATATTTTATACATATTTTAAA 782  
 QY 2218 TAATTTTATGCTGAAACAGAGTTTGGGCAATTTGGACCATTCAGAAAGCAAGTGTAC 2277  
 Db 783 TAATTTTATGCTGAAACAGAGTTTGGGCAATTTGGACCATTCAGAAAGCAAGTGTAC 842  
 QY 2278 TAATTTTATGCTGAAACAGAGTTTGGGCAATTTGGACCATTCAGAAAGCAAGTGTAC 2337  
 Db 843 TAATTTTATGCTGAAACAGAGTTTGGGCAATTTGGACCATTCAGAAAGCAAGTGTAC 902  
 QY 2338 TGATCCGAGTACTTTGGAGCCAGAGTGTGATCTTTGAGCCAGGAGTTTGAGGC 2397  
 Db 903 TGATCCGAGTACTTTGGAGCCAGAGTGTGATCTTTGAGCCAGGAGTTTGAGGC 962  
 QY 2398 CAGACTGCAACACAGTGTGAGTCTGTTTCTCAAAATAATTTAAATAATTTAGCAGGTGT 2457  
 Db 963 CAGACTGCAACACAGTGTGAGTCTGTTTCTCAAAATAATTTAAATAATTTAGCAGGTGT 1022  
 QY 2458 GGTGGTGCACACCTGTAGTCCAGGTACTCAGAGGCTGAGG-TAGTAGGATTTTGGAG 2516  
 Db 1023 GGTGGTGCACACCTGTAGTCCAGGTACTCAGAGGCTGAGG-TAGTAGGATTTTGGAG 1082  
 QY 2517 ACTGGGAGGTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 2560  
 Db 1083 ACTGGGAGGTGAGGCTGAACTGAGCCAGGATCTTGCCACCACA 1126

RESULT 9

ABV26653

ID ABV26653 standard; cdna; 1128 BP.

XX

AC ABV26653;

XX

DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cdna 26644.  
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX OS Homo sapiens.  
 XX FN WO200160860-A2.  
 XX PD 23-AUG-2001.  
 XX PP 20-FEB-2001; 2001WO-US005171.  
 XX PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA Schlegel R, Endege WO, Monahan JE;  
 XX WPI; 2001-662795/76.  
 DR Novel isolated nucleic acid molecule associated with cancerous state of  
 XX prostate cells and correlating with presence of prostate cancer, useful  
 XX for detecting presence of prostate cancer, stage of prostate cancer.  
 XX Claim 1; Page 5383; 11750pp; English.  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;  
 Query Match 28.4%; Score 1068; DB 5; Length 1128;  
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;  
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;  
 QY 1438 CCAATTTTCAAATGCCAACATCTCCCATCCCAATTACCTATTTTCATCTTTGAGGTGAAT 1497  
 Db 4 CTACAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTCACTTTGAGGGTAAT 62  
 QY 1498 CTACTCAATAAACTGTGTAAAGACGACGACGACGACGACGACGACGACGACGACGACG 1557  
 Db 63 CTACTCAATAAACTGTGTAAAGACGACGACGACGACGACGACGACGACGACGACGACG 122  
 QY 1558 AATTTTCTTTTCTATGTACTGGATATTTTGCATATAAACTGTCAGTAATAGTTCAAA 1617  
 Db 123 AATTTTCTTTTCTATGTACTGGATATTTTGCATATAAACTGTCAGTAATAGTTCAAA 182  
 QY 1618 AATTAATAGTTTTTGCATTTGGCTTTCTGAGAAGAGAAATGAAAGTGTCAAAAATAA 1677  
 Db 183 AATTAATAGTTTTTGCATTTGGCTTTCTGAGAAGAGAAATGAAAGTGTCAAAAATAA 242  
 QY 1678 AAAAGATGAATGAAGCATATATAATGTCAATTTTTTCAATTTTCTAGTCAACAGAGA 1737  
 Db 243 AAAAGATGAATGAAGCATATATAATGTCAATTTTTTCAATTTTCTAGTCAACAGAGA 302  
 QY 1738 ATCGAAGATTCTGTTCAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATG 1797





CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;  
 Query Match 28.4%; Score 1068; DB 5; Length 1128;  
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;  
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;  
 QY 1438 CCAAAATTTCAAATGCCAATCTCCCATCCCAATCTCTGAGGTGTAAT 1497  
 DB 4 CTACAAATTTAAATGCAACATCTCCCATCCCAATCTCTGAGGTGTAAT 62  
 QY 1498 CTACTCAATAAATCTGTAGACAGTGCAGACAGCCCTTCTTAACCTGACATTTACTTC 1557  
 DB 63 CTACTCAATAAATCTGTAGACAGTGCAGACAGCCCTTCTTAACCTGACATTTACTTC 122  
 QY 1558 AATTTTCTTTCTTCTATGTACTGATATTTTGCATATATAAATCTGAGTAACTGTCAAA 1617  
 DB 123 AATTTTCTTTCTTCTATGTACTGATATTTTGCATATATAAATCTGAGTAACTGTCAAA 182  
 QY 1618 AATTAATAGTTTTTGACATTTGCTTTCTGAGAGAGAAATTTGAAAGTGTACAAAATAA 1677  
 DB 183 AATTAATAGTTTTTGACATTTGCTTTCTGAGAGAGAAATTTGAAAGTGTACAAAATAA 242  
 QY 1678 AAAAAGATGAATGAAGCATATATAATTTGCAATTTTCTCAATTTCTGAGTCAACAGAG 1737  
 DB 243 AAAAAGATGAATGAAGCATATATAATTTGCAATTTTCTCAATTTCTGAGTCAACAGAG 302  
 QY 1738 ATCGAAGGATCTTCTTCAATATATTAGTAAATTTGAAATTTGAAATTTGCTTATTTTG 1797  
 DB 303 ATCGAAGGATCTTCTTCAATATATTAGTAAATTTGAAATTTGAAATTTGCTTATTTTG 362  
 QY 1798 TTGCAACACACTAGTTTAAATTTAACTGTGACTAGTTTATCTTACCGAAGGTGATGTGT 1857  
 DB 363 TTGCAACACACTAGTTTAAATTTAACTGTGACTAGTTTATCTTACCGAAGGTGATGTGT 422  
 QY 1858 AGTTTCTGTTTTTAAATTTCAAGCAATCTGAAATTAATCCATTAATTTGCTTTCTTT 1917  
 DB 423 AGTTTCTGTTTTTAAATTTCAAGCAATCTGAAATTAATCCATTAATTTGCTTTCTTT 462  
 QY 1918 CCCAAGAAGTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGCAAAAGCTTAATGA 1977  
 DB 483 CCCAAGAAGTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGCAAAAGCTTAATGA 542  
 QY 1978 CAATGCATTCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 2037  
 DB 543 CAATGCATTCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 602  
 QY 2038 ATGAGATGCTTGGACAGAGTGTGTTTGGATTTTCAATTTTATTTTGGATTTTGGAA 2097  
 DB 603 ATGAGATGCTTGGACAGAGTGTGTTTGGATTTTCAATTTTATTTTGGATTTTGGAA 662  
 QY 2098 TTTCCATACATATAATGAGAGAGTTGGAATGGAATTCAGTCTAATCAATAAATTCAC 2157  
 DB 663 TTTCCATACATATAATGAGAGAGTTGGAATGGAATTCAGTCTAATCAATAAATTCAC 722  
 QY 2158 TTATGTTTGTATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATTTTAAA 2217  
 DB 723 TTATGTTTGTATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATTTTAAA 782  
 QY 2218 TAATTTTATGCTGGAACAGATTTTGGACATTTGGACCATCAGAAAGCAGAGTGTAC 2277  
 DB 783 TAATTTTATGCTGGAACAGATTTTGGACATTTGGACCATCAGAAAGCAGAGTGTAC 842  
 QY 2278 TAATTTTATGCTGGAACAGATTTTGGACATTTGGACCATCAGAAAGCAGAGTGTAC 2337  
 DB 843 TAATTTTATGCTGGAACAGATTTTGGACATTTGGACCATCAGAAAGCAGAGTGTAC 902  
 QY 2338 TGATCCAGTACTTTGGAGCCNAGACAGGTGATCTTTGAGCCAGAGTTTGAAGC 2397  
 DB 903 TGATCCAGTACTTTGGAGCCNAGACAGGTGATCTTTGAGCCAGAGTTTGAAGC 962

QY 2398 CAGACTGCACAAACACAGTGCAGCTGTTTCTCAAAATTAATTAATAATTTAGCAGGTGT 2457  
 DB 963 CAGACTGCACAAACACAGTGCAGCTGTTTCTCAAAATTAATTAATAATTTAGCAGGTGT 1022  
 QY 2458 GGTGGTGCACACCTGTAGTCCAGGCTACTCAGAGGCTGAGG-TAGTAGGATTTGTTGAG 2516  
 DB 1023 GGTGGTGCACACCTGTAGTCCAGGCTACTCAGAGGCTGAGGTTATTTGGATTTGTTGAG 1082  
 QY 2517 ACTGGAGGTTGAGGCTGAACTCAGCCAGGATCTTGCCACCACA 2560  
 DB 1083 ACTGGAGGTTGAGGCTGAACTCAGCCAGGATCTTGCCACCACA 1126  
 RESULT 7  
 ID ABV20439 standard; cdna; 1128 BP.  
 XX AC ABV20439;  
 XX AC ABV20439;  
 DT 13-SEP-2002 (first entry)  
 XX Human prostate expression marker cdna 20430.  
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 OS WO2001160860-A2.  
 PN 23-AUG-2001.  
 PD 20-FEB-2001; 2001WO-US005171.  
 PF 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 03-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX Schlegel R, Endege WO, Monahan JE;  
 XX WPI; 2001-662795/76.  
 DR Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX Claim 1; Page 3347; 11750pp; English.  
 PS The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;  
 Query Match 28.4%; Score 1068; DB 5; Length 1128;  
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;  
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;  
 QY 1438 CCAAAATTTCAAATGCCAATCTCCCATCCCAATCTCTGAGGTGTAAT 1497

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 28.4%; Score 1068; DB 5; Length 1128;  
Best Local Similarity 98.0%; Pred. No. 1.7e-144;  
Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 1438 CCAAAATTTCAATGCGACATCTCCCATCCCAATTACCTATTTCATCTTTGAGGTGAAT 1497  
DB 4 CTACAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTCATCTTTGAGGTGAAT 62  
QY 1498 CTACTCAATAAATCTGTGAAGACCACTGACAGACCCCTTTGCTTAACCTGACATTTTACTTC 1557  
DB 63 CTACTCAATAAATCTGTGAAGACCACTGACAGACCCCTTTGCTTAACCTGACATTTTACTTC 122  
QY 1558 AATTTTCTTTCTATGACTGATATTTTGCATATAAATCTGAGTAATAGTTCAA 1617  
DB 123 AATTTTCTTTCTATGACTGATATTTTGCATATAAATCTGAGTAATAGTTCAA 182  
QY 1618 AATTAATAGTTTTCGATCTGCTTTCTGAGAGAGAAATTCAGAGTGTCAAAATAA 1677  
DB 183 AATTAATAGTTTTCGATCTGCTTTCTGAGAGAGAAATTCAGAGTGTCAAAATAA 242  
QY 1678 AAAAAAGATGAAGACATATAATTTCTCAATTTTCTAGTCAACACAGAGA 1737  
DB 243 AAAAAAGATGAAGACATATAATTTCTCAATTTTCTAGTCAACACAGAGA 302  
QY 1738 ATCGAAGGATCTGTTCAATATTTAGTAAAAATTTGAAATTAACCTGTGCTTAATTTTG 1797  
DB 303 ATCGAAGGATCTGTTCAATATTTAGTAAAAATTTGAAATTAACCTGTGCTTAATTTTG 362  
QY 1798 TTTGCAACACACTAGTTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGT 1857  
DB 363 TTTGCAACACACTAGTTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGT 422  
QY 1858 AGTTTCTGTTTTTAAAAATTCAGCAAACTGGAAAAATTAATCCATCTAATTTATGCTTTT 1917  
DB 423 AGTTTCTGTTTTTAAAAATTCAGCAAACTGGAAAAATTAATCCATCTAATTTATGCTTTT 482  
QY 1918 CCCAAGAGTTTAAATGATATGCGAGCTTCCCTAATTTGGAGACAAAAGCCTTAATGA 1977  
DB 483 CCCAAGAGTTTAAATGATATGCGAGCTTCCCTAATTTGGAGACAAAAGCCTTAATGA 542  
QY 1978 CAATGATTCATATATATTTTCTAGTATACAGATACAGATGAGTATCCCTTAG 2037  
DB 543 CAATGATTCATATATATTTTCTAGTATACAGATACAGATGAGTATCCCTTAG 602  
QY 2038 ATGAGATGCTTGGGACCAAGGTGTTTGGATTTTCAGATTTAATTTTGGATTTTGAATA 2097  
DB 603 ATGAGATGCTTGGGACCAAGGTGTTTGGATTTTCAGATTTAATTTTGGATTTTGAATA 662  
QY 2098 TTTCCATATATATAGAGAGTTGGAAAAATGGGATTCAGTCTAATCAATAAATTCAC 2157  
DB 663 TTTCCATATATATAGAGAGTTGGAAAAATGGGATTCAGTCTAATCAATAAATTCAC 722  
QY 2158 TTATGTTTGGATATACACCTTATCTGAATAGCCTGAAGGTAAATTTTATACATAATTTTAA 2217  
DB 723 TTATGTTTGGATATACACCTTATCTGAATAGCCTGAAGGTAAATTTTATACATAATTTTAA 782  
QY 2218 TAATTTTATGCTTGAACACAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2277  
DB 1128 TAATTTTATGCTTGAACACAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 842

DB 783 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 842  
QY 2278 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2337  
DB 843 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 902  
QY 2338 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2397  
DB 903 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 962  
QY 2398 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2457  
DB 963 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 1022  
QY 2458 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2516  
DB 1023 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 1082  
QY 2517 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2560  
DB 1083 TAAATTTTATGCTTGAACACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 1126

RESULT 6  
ABV20806  
ID ABV20806 standard; cDNA; 1128 BP.  
XX AC ABV20806;  
XX DT 13-SEP-2002 (first entry)  
XX DE Human prostate expression marker cDNA 20797.  
XX EX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW Pharmacogenomic marker; gene; ss.  
XX OS Homo sapiens.  
XX PN WO200160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US005171.  
XX PR 17-FEB-2000; 2000US-0183319P.  
XX PR 16-MAR-2000; 2000US-0189862P.  
XX PR 25-MAY-2000; 2000US-0207454P.  
XX PR 09-JUN-2000; 2000US-0211314P.  
XX PR 18-JUL-2000; 2000US-0219007P.  
XX PR 13-DEC-2000; 2000US-0255281P.  
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PI Schlegel R, Endege WO, Monahan JB;  
XX PI WPI; 2001-662795/76.  
XX DR Novel isolated nucleic acid molecule associated with cancerous state of  
XX PT prostate cells and correlating with presence of prostate cancer, useful  
XX PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX PS Claim 1; Page 3412-3413; 11750pp; English.  
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising  
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX CC specification or its complement. (I) is useful for: (a) assessing whether  
XX CC a patient is afflicted with prostate cancer; (b) monitoring the  
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy  
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
XX CC determining whether prostate cancer has metastasized in a patient; (h)

PI	Schlegel R, Endege WO, Monahan JE;	Qy	2098	TTTCCATACATATATGAGAGAGTTGGAAAATGGGATTCAAGTCTAATCATATAAATTCAC	2157
XX		Db	663	TTTCCATACATATATGAGAGAGTTGGAAAATGGGATTCAAGTCTAATCATATAAATTCAC	722
XX	WPI; 2001-662795/76.	Qy	2158	TTATGTTTGTATATACACCTTATCTGAATAGCTGGAAGGTAATTTTATACAATATTTTAAA	2217
PT	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.	Db	723	TTATGTTTGTATATACACCTTATCTGAATAGCTGGAAGGTAATTTTATACAATATTTTAAA	782
XX		Qy	2218	TAATTTTATGCTGAAACACAGAGTTTGGCACCATTGGACCATCAGAAAACGAGAAGTGTAC	2277
XX	Claim 1; Page 5346-5347; 11750pp; English.	Db	783	TAATTTTATGCTGAAACACAGAGTTTGGCACCATTGGACCATCAGAAAACGAGAAGTGTAC	842
CC	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker	Qy	2278	TATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGATCGCAGTTTGAAGC	2337
CC		Db	843	TATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGATCGCAGTTTGAAGC	902
CC		Qy	2338	TGATCCGAGTACTTTGGGAAGCCAAAGACAGTGTGATCTTTGAGCCGAGGTTTGAAGC	2397
CC		Db	903	TGATCCGAGTACTTTGGGAAGCCAAAGACAGTGTGATCTTTGAGCCGAGGTTTGAAGC	962
CC		Qy	2398	CAGACTGCAACAACACAGTGTGAGTCTGTTTCTACAAATAATTAATAAATTTAGCCAGTGT	2457
CC		Db	963	CAGACTGCAACAACACAGTGTGAGTCTGTTTCTACAAATAATTAATAAATTTAGCCAGTGT	1022
Qy	Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;	Qy	2458	GGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGG-TAGTAGGATTTGTGAG	2516
Best Match	28.4%; Score 1068; DB 5; Length 1128;	Db	1023	GGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGG-TAGTAGGATTTGTGAG	1082
Best Local Similarity	98.0%; Pred. No. 1.7e-144;	Qy	2517	ACTGGGAGGTTTGGAGGCTGAACTGAGCCAGGATCTTGCCACCACA	2560
Matches 1102; Conservative	0; Mismatches 20; Indels 2; Gaps 2;	Db	1083	ACTGGGAGGTTTGGAGGCTGAACTGAGCCAGGATCTTGCCACCACA	1126
Qy	1438 CCAAAATTTCAAAATGCCAACATCTCCCATCCAAATACCTATTTTCATCTTTGAGGTGAAT	RESULT 5			
Db	4 CTACAAATTTAAATGCAACATCTGCCATCAA-TACCTATTTTCATCTTTGAGGGTAAAT	ABV20617			
Qy	1498 CTACTCAATAACTGTAGACACAGTACAGACCCCTTGTCTAATCTGATTAATCTTCTC	ID	ABV20617	standard; cdna; 1128 BP.	
Db	63 CTACTCAATAACTGTAGACACAGTACAGACCCCTTGTCTAATCTGATTAATCTTCTC	XX	AC	ABV20617;	
Qy	1558 AAATTTTCTTTTCTATGTAATTTTCTGATATAAATTTGATTAATTTGATTAATTTG	XX	DT	13-SEP-2002 (first entry)	
Db	123 AAATTTTCTTTTCTATGTAATTTTCTGATATAAATTTGATTAATTTGATTAATTTG	DE	DE	Human prostate expression marker cdna 20608.	
Qy	1618 AATTAATGATTTTGAATTTGCTTTTCTGAGAGAGAAATTTGAAAGTGTACAAAATAA	XX	XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
Db	183 AATTAATGATTTTGAATTTGCTTTTCTGAGAGAGAAATTTGAAAGTGTACAAAATAA	XX	XX	pharmacogenomic marker; gene; ss.	
Qy	1678 AAAAAATGAAATGAAGCATATATAATTTGTCATTTTCTCAATTTTCTAGCCACAGAG	OS	OS	Homo sapiens.	
Db	243 AAAAAATGAAATGAAGCATATATAATTTGTCATTTTCTCAATTTTCTAGCCACAGAG	PN	PN	WO200160860-A2.	
Qy	1738 ATCGAAGGATTTCTTCAAAATATTAGTAAATTTGAAATTTGAAATTTGAAATTTG	XX	XX	23-AUG-2001.	
Db	303 ATCGAAGGATTTCTTCAAAATATTAGTAAATTTGAAATTTGAAATTTGAAATTTG	PF	PF	20-FEB-2001; 2001WO-US005171.	
Qy	1798 TTGCAACACACTAGTTAAATTTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTG	XX	XX	17-FEB-2000; 2000US-0183319P.	
Db	363 TTGCAACACACTAGTTAAATTTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTG	PR	PR	16-MAR-2000; 2000US-0189862P.	
Qy	1858 AGTTCTGGTTTAAATTTCAAGCAACTGGAAATTAATCCATTAATTAATGCTTTCTTT	PR	PR	25-MAY-2000; 2000US-0207454P.	
Db	423 AGTTCTGGTTTAAATTTCAAGCAACTGGAAATTAATCCATTAATTAATGCTTTCTTT	PR	PR	09-JUN-2000; 2000US-0211314P.	
Qy	1918 CCCAAGAAGTTTATTAATGATATGATGCTTCTTAAATTTGGAGCAAAAGCCTTAATGA	PR	PR	18-JUL-2000; 2000US-0219007P.	
Db	483 CCCAAGAAGTTTATTAATGATATGATGCTTCTTAAATTTGGAGCAAAAGCCTTAATGA	PR	PR	13-DEC-2000; 2000US-0255281P.	
Qy	1978 CAATGCAATCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG	PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
Db	543 CAATGCAATCATATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG	XX	Schlegel R, Endege WO, Monahan JE;		
Qy	2038 ATGAGATGCTTGGACACAGAGTGTGTTTGGATTTTCAATTTTGGATTTTGGAAATA	XX	WPI; 2001-662795/76.		
Db	603 ATGAGATGCTTGGACACAGAGTGTGTTTGGATTTTCAATTTTGGATTTTGGAAATA	XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.		
		PS	Claim 1; Page 3379; 11750pp; English.		

XX PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0198862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA Schlegel R, Endege WO, Monahan JE;  
 PI WPI; 2001-662795/76.  
 DR Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX Claim 1; Page 5285-5286; 11750pp; English.  
 PS The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;  
 SQ Query Match 28.4%; Score 1068; DB 5; Length 1128;  
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;  
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;  
 QY 1438 CCAAAATTTCAAAATGCCAAATCTCCCATCCCAATTAACCTATTTTCATCTTTGAGGTGAAT 1497  
 DB 4 CTACAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTTCATCTTTGAGGTGAAT 62  
 QY 1498 CTACTCAATAAAGTGTGAGACCTGACGACCCCTTTGCTAACTGACATTTACTTC 1557  
 DB 63 CTACTCAATAAAGTGTGAGACCTGACGACCCCTTTGCTAACTGACATTTACTTC 122  
 QY 1558 AATTTTCTTTCTATGTAAGTATATTTTGCATATAAACTTGCAAGTAAGTTCAAA 1617  
 DB 123 AATTTTCTTTCTATGTAAGTATATTTTGCATATAAACTTGCAAGTAAGTTCAAA 182  
 QY 1618 AATTAATAGTTTGTGACATGCTTTCTGAGAGAGAAATGAAAGTGTCAAAATAA 1677  
 DB 183 AATTAATAGTTTGTGACATGCTTTCTGAGAGAGAAATGAAAGTGTCAAAATAA 242  
 QY 1678 AAAAAAGTGAATGAGACATATATATGTCATTTTCTCAATTTTCTAGTCAACAGAGA 1737  
 DB 243 AAAAAAGTGAATGAGACATATATATGTCATTTTCTCAATTTTCTAGTCAACAGAGA 302  
 QY 1738 ATCGAAGGATTTCTGTTCAATATATTAGTAAAAATTTGAAAAATTTGCTTATATTTTG 1797  
 DB 303 ATCGAAGGATTTCTGTTCAATATATTAGTAAAAATTTGAAAAATTTGCTTATATTTTG 362  
 QY 1798 TTGCAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCAAGGTGATGTGT 1857  
 DB 363 TTGCAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCAAGGTGATGTGT 422  
 QY 1858 AGTTTCTGTTTAAAAATTCAGCAAACTGGAATAAATCCATCTAATATGCTTCTTT 1917  
 DB 423 AGTTTCTGTTTAAAAATTCAGCAAACTGGAATAAATCCATCTAATATGCTTCTTT 482  
 QY 1918 CCAAGAAGTTTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATGA 1977

DB 483 CCCAAGAAAGTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAAGCCTTAATGA 542  
 QY 1978 CAATGCATTCATTATATATTTTTTTTGTATAGTTCACAGTATACAGTTGAGTATCCCTTAG 2037  
 DB 543 CAATGCATTCATTATATATTTTTTTTGTATAGTTCACAGTATACAGTTGAGTATCCCTTAG 602  
 QY 2038 ATGAGATGCTTGGGACCCAGAGAGTGTGTTTGGATTTTTCAGATTTATTTTGGATTTTGAATA 2097  
 DB 603 ATGAGATGCTTGGGACCCAGAGAGTGTGTTTGGATTTTTCAGATTTATTTTGGATTTTGAATA 662  
 QY 2098 TTTCCATACATATAATAGAGAGAGTTGGAAAATGGGATTCAGTCTAAATCAATAAATTCAC 2157  
 DB 663 TTTCCATACATATAATAGAGAGAGTTGGAAAATGGGATTCAGTCTAAATCAATAAATTCAC 722  
 QY 2158 TTATGCTTTGATATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATATTTTAA 2217  
 DB 723 TTATGCTTTGATATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATATTTTAA 782  
 QY 2218 TAAATTTTATGCTGAAACAGAGTGTGCGCACATTTGGACCATCAGAAAGCAGAGTGTCTAC 2277  
 DB 783 TAAATTTTATGCTGAAACAGAGTGTGCGCACATTTGGACCATCAGAAAGCAGAGTGTCTAC 842  
 QY 2278 TATTTCAAGTCAAGTGTCTCAAAAAGTTTCAGATTTTAAAGCTGTGATGCGAGTTTCATGCCAG 2337  
 DB 843 TATTTCAAGTCAAGTGTCTCAAAAAGTTTCAGATTTTAAAGCTGTGATGCGAGTTTCATGCCAG 902  
 QY 2338 TGATCCGAGTACTTTGGGAAAGCCAGACAGAGTGGATCTCTTGAGCCCGAGGAGTTTGAGGC 2397  
 DB 903 TGATCCGAGTACTTTGGGAAAGCCAGACAGAGTGGATCTCTTGAGCCCGAGGAGTTTGAGGC 962  
 QY 2398 CAGACTGCAACACACAGAGTGAACCTCGTTTCTACAAATTAATTAATAAATTTAGCCAGGTGT 2457  
 DB 963 CAGACTGCAACACACAGAGTGAACCTCGTTTCTCAAAATTAATTAATAAATTTAGCCAGGTGT 1022  
 QY 2458 GGTGTGTGACACCTGTAGTCCAGGTAAGTCTGAGGAGGCTGAGG-TAGTAGATTTGTTTGG 2516  
 DB 1023 GGTGTGTGACACCTGTAGTCCAGGTAAGTCTGAGGAGGCTGAGG-TAGTAGATTTGTTTGG 1082  
 QY 2517 ACTGGGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCACA 2560  
 DB 1083 ACTGGGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCACA 1126  
 XX RESULT 4  
 XX ABV26458  
 XX ID ABV26458 standard; cDNA; 1128 BP.  
 XX AC ABV26458;  
 XX DT 16-SEP-2002 (first entry)  
 XX DE Human prostate expression marker cDNA 26449.  
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 XX KW pharmacogenomic marker; gene; ss.  
 XX OS Homo sapiens.  
 XX XX WO200160860-A2.  
 XX PD 23-AUG-2001.  
 XX PF 20-FEB-2001; 2001WO-US005171.  
 XX PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA  
 XX

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 3320-3321; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 28.4%; Score 1068; DB 5; Length 1128;

Best Local Similarity 98.0%; Pred. No. 1.7e-144;

Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

1438 CCAAAATTCAAATGCCAACATCTCCCATCCCAATTAACCTATTTTCATCTTTGAGGTGTAAT 1497

4 CTACAAATTTAATGACACATCTGCCCATCCAA-TACCTATTTCACTTTTGGGGGTAAT 62

1498 CTACTCAATAAATGTGTAGAACACAGTACAGACCCCTTTGCTAACTGACATTTACTTC 1557

63 CTACTCAATAAATGTGTAGAACACAGTACAGACCCCTTTGCTAACTGACATTTACTTC 122

1558 AATTTTCTTTTCTATGTAAGTATTTTTCATATAAATTTGAGTAATAGTTCAAA 1617

123 AATTTTCTTTTCTATGTAAGTATTTTTCATATAAATTTGAGTAATAGTTCAAA 182

1618 AATTAATAGTTTTTGTACATTTGCTTTCTGAGAGAGAAATTCGAAAGTGTCAAAAATAA 1677

183 AATTAATAGTTTTTGTACATTTGCTTTCTGAGAGAGAAATTCGAAAGTGTCAAAAATAA 242

1678 AAAAAGATGAATGAAGCATATATAATTTTCAATTTTCTAGTCAACAGAGA 1737

243 AAAAAGATGAATGAAGCATATATAATTTTCAATTTTCTAGTCAACAGAGA 302

1738 ATCGAAGGATTCGTTCAATATATAGTAAATTTGAAATTAATCTGTGCTTATATTTG 1797

303 ATCGAAGGATTCGTTCAATATATAGTAAATTTGAAATTAATCTGTGCTTATATTTG 362

Qy	1798	TTTGCACACACACTAGTTAAATTTAACTGTGCTAGTTATCTCTACCGAAGGTGATGTGT	1857
Db	363	TTTGCACACACACTAGTTAAATTTAACTGTGCTAGTTATCTCTACCGAAGGTGATGTGT	422
Qy	1858	AGTTTCTGGTTTAAATTTCAAGCAACTGCGAAATTAATCCATCTAATTAATGCTTTCTTT	1917
Db	423	AGTTTCTGGTTTAAATTTCAAGCAACTGCGAAATTAATCCATCTAATTAATGCTTTCTTT	482
Qy	1918	CCCAAGAAAGTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAGCCTTAATTTGA	1977
Db	483	CCCAAGAAAGTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAGCCTTAATTTGA	542
Qy	1978	CAATGCATTCATTATATATTTTTTTGATAGTATACAGTATACGAGTTGAGTATCCCTTAG	2037
Db	543	CAATGCATTCATTATATATTTTTTTGATAGTATACAGTATACGAGTTGAGTATCCCTTAG	602
Qy	2038	ATGAGATGCTTTGGGACCGAGAGTGTTCGATTTTCAGATTTATTTTGGATTTTGGATA	2097
Db	603	ATGAGATGCTTTGGGACCGAGAGTGTTCGATTTTCAGATTTATTTTGGATTTTGGATA	662
Qy	2098	TTTCCATACATATAATGAGAGAGTTTGGAAATGGGATTTCAAGTCTAATCATATAAATTCAC	2157
Db	663	TTTCCATACATATAATGAGAGAGTTTGGAAATGGGATTTCAAGTCTAATCATATAAATTCAC	722
Qy	2158	TTATGTTTGTATATACACCTTATCTGAATAGCCTGAAGGTAATTTTATACAAATTTTAA	2217
Db	723	TTATGTTTGTATATACACCTTATCTGAATAGCCTGAAGGTAATTTTATACAAATTTTAA	782
Qy	2218	TAATTTTATGCTGAAACAGAGTTTGGCACATTTGGACCATTCAGAAAGCAGAGAGTGTAC	2277
Db	783	TAATTTTATGCTGAAACAGAGTTTGGCACATTTGGACCATTCAGAAAGCAGAGAGTGTAC	842
Qy	2278	TATTTCAAGTCAAGTGTCTCAAAAAGTTTCAAGTGTGATGCTGATGCTCATGCCAG	2337
Db	843	TATTTCAAGTCAAGTGTCTCAAAAAGTTTCAAGTGTGATGCTGATGCTCATGCCAG	902
Qy	2338	TGATCCGAGTACTTTGGGAAGCCCAAGCAGGTGGATCTCTTGAGCCCGAGAGTTTGAAGC	2397
Db	903	TGATCCGAGTACTTTGGGAAGCCCAAGCAGGTGGATCTCTTGAGCCCGAGAGTTTGAAGC	962
Qy	2398	CAGACTGCACACACACAGTGCAGCTCTGTTTCTACAAATTAATTAATAAATTTAGCCAGGT	2457
Db	963	CAGACTGCACACACACAGTGCAGCTCTGTTTCTACAAATTAATTAATAAATTTAGCCAGGT	1022
Qy	2458	GGTGTGCACACCTGTAGTCCCAGGTACTCAGAGGCTGAGG-TAGTAGGATTTGTTTGG	2516
Db	1023	GGTGTGCACACCTGTAGTCCCAGGTACTCAGAGGCTGAGGTTATTTGGATTTGTTTGG	1082
Qy	2517	ACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTCCACACCA 2560	
Db	1083	ACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTCCACACCA 1126	
RESULT 3			
ABV26130			
ID	ABV26130	standard; cDNA; 1128 BP.	
XX	ABV26130;		
AC	ABV26130;		
XX	XX		
DT	16-SEP-2002	(first entry)	
XX	XX		
DE	Human prostate expression marker cDNA 26121.		
XX	XX		
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;		
KW	pharmacogenomic marker; gene; ss.		
OS	Homo sapiens.		
XX	XX		
PN	WO200160860-A2.		
XX	XX		
PD	23-AUG-2001.		
XX	XX		
PF	20-FEB-2001; 2001WO-US005171.		

1741 GAAGGATCTGTTCAAAATATTAGTAAATAATGAAATAAATGAGCTTATATTGTTT 1800  
 1801 GCAACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860  
 1801 GCAACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860  
 1861 TTCTGGTTTAAATTAACCTGGAATAATGAAATAATGAAATAATGAAATAATGAAATAAT 1920  
 1861 TTCTGGTTTAAATTAACCTGGAATAATGAAATAATGAAATAATGAAATAATGAAATAAT 1920  
 1921 AAGAGGTTTAAATTAACCTGGAATAATGAAATAATGAAATAATGAAATAATGAAATAAT 1980  
 1921 AAGAGGTTTAAATTAACCTGGAATAATGAAATAATGAAATAATGAAATAATGAAATAAT 1980  
 1981 TGCAATTCATTAAT 2040  
 1981 TGCAATTCATTAAT 2040  
 2041 AGATGCTTGGACCAAGAGTGTGTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2100  
 2041 AGATGCTTGGACCAAGAGTGTGTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2100  
 2101 CCATACATATATATGAGAGTGTGTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2160  
 2101 CCATACATATATGAGAGTGTGTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2160  
 2161 TGTTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACAAATTTTAAATAA 2220  
 2161 TGTTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACAAATTTTAAATAA 2220  
 2221 TTTTATGCTGAAACAGAGTGTGTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2280  
 2221 TTTTATGCTGAAACAGAGTGTGTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTT 2280  
 2281 TTCAAGTCACTGCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGAGTTTCAAGTGTGA 2340  
 2281 TTCAAGTCACTGCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGAGTTTCAAGTGTGA 2340  
 2341 TCCGAGTACTTTGGGAAGCCAAAGACAGAGTGGATCTCTTGAGCCAGGAGTTTGGGCCAG 2400  
 2341 TCCGAGTACTTTGGGAAGCCAAAGACAGAGTGGATCTCTTGAGCCAGGAGTTTGGGCCAG 2400  
 2401 ACTGCACAAACAGTGTGAGACCTGTTTCTACAAATAATTAATAAATAGCAGGTGTGTT 2460  
 2401 ACTGCACAAACAGTGTGAGACCTGTTTCTACAAATAATTAATAAATAGCAGGTGTGTT 2460  
 2461 GGTGCACACCTGTAGTCCAGGTACTCAGAGGCTGAGGTAGTGTAGTGTGTTGAGACTG 2520  
 2461 GGTGCACACCTGTAGTCCAGGTACTCAGAGGCTGAGGTAGTGTAGTGTGTTGAGACTG 2520  
 2521 GGAGGTGAGGCTGAACTGAGCCAGGATCTTGCCACCAATTCAGCTTTGGGCCAACAGAG 2580  
 2521 GGAGGTGAGGCTGAACTGAGCCAGGATCTTGCCACCAATTCAGCTTTGGGCCAACAGAG 2580  
 2581 TGAGACCTGCTCAAAAAGTTTCAAGTTTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2640  
 2581 TGAGACCTGCTCAAAAAGTTTCAAGTTTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2640  
 2641 ATTAGGGATTTTCAACCTGTACTGACCTTTTGTAGTCAATGACAGCAATTAATCAATAGG 2700  
 2641 ATTAGGGATTTTCAACCTGTACTGACCTTTTGTAGTCAATGACAGCAATTAATCAATAGG 2700  
 2701 GACTCCAGATAAATCAATTTGCTGTATACAAATTTTGGCTCTCTATTTCAACGAAATCTTAT 2760  
 2701 GACTCCAGATAAATCAATTTGCTGTATACAAATTTTGGCTCTCTATTTCAACGAAATCTTAT 2760  
 2761 GGCCTCTTGGTGTATTTAATGTCGGAGGGAACCAATAGAAATTTTGGCAATCTTCA 2820  
 2761 GGCCTCTTGGTGTATTTAATGTCGGAGGGAACCAATAGAAATTTTGGCAATCTTCA 2820  
 2821 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 2880

2821 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 2880  
 2881 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 2880  
 2881 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 2880  
 2881 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 2880  
 2941 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 3000  
 2941 AAAGTCATTTCTGTCAAAATATGTCAGTCTGTGTAGATATTAGCCAAATTTTAGGAAATGAC 3000  
 3001 GGTAAATAAAACAGTATTGATTTGGTAGAAGAACGTTGAAATCCAGAGACATCAATGTCTT 3060  
 3001 GGTAAATAAAACAGTATTGATTTGGTAGAAGAACGTTGAAATCCAGAGACATCAATGTCTT 3060  
 3061 CTGGTGGTTTCAACATTAAGCCACAGAGATGTCCTTAATCTTTCCGAGATCTAGTTTTCAG 3120  
 3061 CTGGTGGTTTCAACATTAAGCCACAGAGATGTCCTTAATCTTTCCGAGATCTAGTTTTCAG 3120  
 3121 CAAGACAGGATTTAAGAAATGTAATCTTATGTTGGTTTATGAAGAACCAATAGAATCAAT 3180  
 3121 CAAGACAGGATTTAAGAAATGTAATCTTATGTTGGTTTATGAAGAACCAATAGAATCAAT 3180  
 3181 GCTGTATAGTGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGCAATATAAT 3240  
 3181 GCTGTATAGTGTCTTTTAACTGTAATTTTGTGAAGCTTATCTTTTATGCAATATAAT 3240  
 3241 ATTTGAAACATTTTACATTTGTTTATATATTTTAACTGTAATTTTGTGAAGCTTATATA 3300  
 3241 ATTTGAAACATTTTACATTTGTTTATATATTTTAACTGTAATTTTGTGAAGCTTATATA 3300  
 3301 AAGAAATGTAACCACTGTAAGGGTAGAGTTTATAAGAAATTTTGTCAAATGTAATCACCCA 3360  
 3301 AAGAAATGTAACCACTGTAAGGGTAGAGTTTATAAGAAATTTTGTCAAATGTAATCACCCA 3360  
 3361 TGAGTCACTCTCTTATGAAGACAGAACACAGTACATCTCCAGAGAGTTTCCACAGTG 3420  
 3361 TGAGTCACTCTCTTATGAAGACAGAACACAGTACATCTCCAGAGAGTTTCCACAGTG 3420  
 3421 CTCCTTTTCCCTGAGTTTCAACAGTCTGCGCAACCAATGATCTGCTCGTATAATATAA 3480  
 3421 CTCCTTTTCCCTGAGTTTCAACAGTCTGCGCAACCAATGATCTGCTCGTATAATATAA 3480  
 3481 CTGTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTTATTTTGTGTGTGTAAGGCTT 3540  
 3481 CTGTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTTATTTTGTGTGTGTAAGGCTT 3540  
 3541 CTTTGTAGTCAATTAATAATTTTGTAGATTTTGTAGATTTTGTAGATTTTGTAGATTTTGT 3600  
 3541 CTTTGTAGTCAATTAATAATTTTGTAGATTTTGTAGATTTTGTAGATTTTGTAGATTTTGT 3600  
 3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATATAATTTTGTATCTAATCA 3660  
 3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATATAATTTTGTATCTAATCA 3660  
 3661 CTGATGGATATGAGGATTTTAAAGTTTGTACATTTTGAATTAAGTGTGCTATAATGAA 3720  
 3661 CTGATGGATATGAGGATTTTAAAGTTTGTACATTTTGAATTAAGTGTGCTATAATGAA 3720  
 3721 AA 3762  
 3721 AA 3762

RESULT 2  
 ID ABV20298  
 XX ABV20298 standard; cDNA; 1128 BP.  
 AC ABV20298;  
 XX ABV20298;  
 DT 13-SEP-2002 (first entry)  
 XX Human prostate expression marker cDNA 20289.  
 DE  
 XX



XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),  
 CC corresponding to genes differentially expressed in mast cells following  
 CC activation or in patients with allergic hypersensitivity disease, (I)  
 CC that encodes proteins (ABB7569-ABB7575) (II) or a protein fragment of  
 CC (II) if at least 6 amino acids, (II) is useful for identifying binding  
 CC partners. (I) or (II) is useful for diagnosing or treating a disease  
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,  
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which  
 CC involves determining the level of expression of (I) or (II). A computer  
 CC system, comprising a database containing information identifying the  
 CC expression level in a tissue or at least one mast cell of (I), is useful  
 CC for presenting information to identify the relative expression level of  
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic  
 CC response in a patient. The protein can also serve as a target that  
 CC modulate gene expression or activity and as an antigen to raise  
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying  
 CC agents that modulate expression of the protein or agents, such as  
 CC agonists or antagonists. The agonists or antagonists are useful for  
 CC modulating biological activity and function of (II) and thus are useful  
 CC for alleviating disease conditions such as allergic hypersensitivity,  
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis  
 XX  
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 3762; DB 6; Length 3762;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAGAAACCGAGTCACTGTGAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCTG 60  
 DB 1 GAGAAACCGAGTCACTGTGAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCTG 60  
 QY 61 GAGAGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGAA 120  
 DB 61 GAGAGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGAA 120  
 QY 121 ATGACTACATTTGAAAGAAACCTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCAATC 180  
 DB 121 ATGACTACATTTGAAAGAAACCTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCAATC 180  
 QY 181 ACTTCTAATCAGGAAACCGAGATGCGAGTGGTCTCGAAGAGTGTCTACACTGTCAAT 240  
 DB 181 ACTTCTAATCAGGAAACCGAGATGCGAGTGGTCTCGAAGAGTGTCTACACTGTCAAT 240  
 QY 241 AATCAGATCCCATCAGAGATCTCCCTGAGCTCCCAATGATGATGAGTATGAGAACT 300  
 DB 241 AATCAGATCCCATCAGAGATCTCCCTGAGCTCCCAATGATGATGAGTATGAGAACT 300  
 QY 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGTTCAGAGACAGAAATATGCCCTT 360  
 DB 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGTTCAGAGACAGAAATATGCCCTT 360  
 QY 361 CTTAGGACTCTGTAGTAGGCTTGTCTCGACCCCATCAGCATGATTAATGAAGTTGTG 420  
 DB 361 CTTAGGACTCTGTAGTAGGCTTGTCTCGACCCATCAGCATGATTAATGAAGTTGTG 420  
 QY 421 TTTCCACACTAAATCTCAAGCTGCTTATACCTTCCAGCAATGAAGCAATGAGAA 480  
 DB 421 TTTCCACACTAAATCTCAAGCTGCTTATACCTTCCAGCAATGAAGCAATGAGAA 480  
 QY 481 TAGCAGACTCTGGGAGAGTTGTTCAACCTCAGCAGTGCATGAACATTTCTTCTGCTA 540  
 DB 481 TAGCAGACTCTGGGAGAGTTGTTCAACCTCAGCAGTGCATGAACATTTCTTCTGCTA 540  
 QY 541 AAGTTTGAATAATATCTTATATATATCTTAGGCAACTCTGATATGAGCAATCTCTG 600  
 DB 541 AAGTTTGAATAATATCTTATATATATCTTAGGCAACTCTGATATGAGCAATCTCTG 600  
 QY 601 TGGCTTAGGTAATCATAGAAATTCACAAATGACCTAAATATTTCTATGTTTTTGC 660  
 DB 601 TGGCTTAGGTAATCATAGAAATTCACAAATGACCTAAATATTTCTATGTTTTTGC 660

QY 661 TTGTAAAGTTTGAGGACATGGAGGTGATAAAAAAACTTTCTTAGGCAATAATATGTAATA 720  
 DB 661 TTGTAAAGTTTGAGGACATGGAGGTGATAAAAAAACTTTCTTAGGCAATAATATGTAATA 720  
 QY 721 TGAATAAATATTTCTAATCCCTGACTAATCTGAATGAGACCTCTTCTAGGCCAAAGAGA 780  
 DB 721 TGAATAAATATTTCTAATCCCTGACTAATCTGAATGAGACCTCTTCTAGGCCAAAGAGA 780  
 QY 781 CCTCAGATGAACCTGAAAGACTGAAATCTGGCCATGATAGGAAGGAGGTGAGACACACC 840  
 DB 781 CCTCAGATGAACCTGAAAGACTGAAATCTGGCCATGATAGGAAGGAGGTGAGACACACC 840  
 QY 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGAACAGGATGAGTCAATGACT 900  
 DB 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGAACAGGATGAGTCAATGACT 900  
 QY 901 GATGAATAAGACTGATGTCGCAATAGAGTCCCAATTCCAACTGACTCTGGTGTAGAT 960  
 DB 901 GATGAATAAGACTGATGTCGCAATAGAGTCCCAATTCCAACTGACTCTGGTGTAGAT 960  
 QY 961 CACACACTGTCTGAGGAGTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGTT 1020  
 DB 961 CACACACTGTCTGAGGAGTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGTT 1020  
 QY 1021 TCCAAACCCCTTATTTAGCTTAAAGCAATCTTTCTACTGACTTCTTAAGTCTTTTGA 1080  
 DB 1021 TCCAAACCCCTTATTTAGCTTAAAGCAATCTTTCTACTGACTTCTTAAGTCTTTTGA 1080  
 QY 1081 CAAAGCTTAACTCTTCAACCAATTCGCAATAGAGTCCCAATTCCAACTGACTCTGGTGTAGAT 1140  
 DB 1081 CAAAGCTTAACTCTTCAACCAATTCGCAATAGAGTCCCAATTCCAACTGACTCTGGTGTAGAT 1140  
 QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAAACCGATGTGCACTTTCCA 1200  
 DB 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAAACCGATGTGCACTTTCCA 1200  
 QY 1201 TTTAATGATTTATGCTTGTGTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260  
 DB 1201 TTTAATGATTTATGCTTGTGTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260  
 QY 1261 GTGACCTGACACCTCAGGCACACTTTCTCAGGACCTCTCTGAGGTGTATCCCGAGCCAT 1320  
 DB 1261 GTGACCTGACACCTCAGGCACACTTTCTCAGGACCTCTCTGAGGTGTATCCCGAGCCAT 1320  
 QY 1321 GGTAAAGTCAATGTTGGCTCAGAACTCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380  
 DB 1321 GGTAAAGTCAATGTTGGCTCAGAACTCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380  
 QY 1381 TTACCAATAGTCTCCACAAATATATGTCCAAAGATCTTCAATTCGAAGCTCTCTCACC 1440  
 DB 1381 TTACCAATAGTCTCCACAAATATATGTCCAAAGATCTTCAATTCGAAGCTCTCTCACC 1440  
 QY 1441 AATTTCAAATGCCAACTCTCCCATCTCAATTTACCTTATTTCTTCTTGGGTGTATCTTA 1500  
 DB 1441 AATTTCAAATGCCAACTCTCCCATCTCAATTTACCTTATTTCTTCTTGGGTGTATCTTA 1500  
 QY 1501 CTCATTAATGCTGTGAAGACCAAGTGTGAGACCTTTGCTTAACTGACATTTTCAAT 1560  
 DB 1501 CTCATTAATGCTGTGAAGACCAAGTGTGAGACCTTTGCTTAACTGACATTTTCAAT 1560  
 QY 1561 TTTTCTTTTCTGATGCTGAGATATTTTGGCATATATACTTTCAGTAAATAGTTTCAAAAT 1620  
 DB 1561 TTTTCTTTTCTGATGCTGAGATATTTTGGCATATATACTTTCAGTAAATAGTTTCAAAAT 1620  
 QY 1621 TAATAGTTTTTGCATTTTCTGAGAAGAAATTTGAAAGTGTCAAAAAATAAAAA 1680  
 DB 1621 TAATAGTTTTTGCATTTTCTGAGAAGAAATTTGAAAGTGTCAAAAAATAAAAA 1680  
 QY 1681 AAGATGAATGAAGCATATATAATTTGCTCAATTTTCTCAATTTTCTAGTCAACAGAGAAATC 1740  
 DB 1681 AAGATGAATGAAGCATATATAATTTGCTCAATTTTCTCAATTTTCTAGTCAACAGAGAAATC 1740  
 QY 1741 GAAGGATTTCTGTTCAATATAGTAAATAATTCGAAATAAACTTGTGCTTATATTTGTTT 1800

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 12:17:23 ; Search time 1654 Seconds  
(without alignments)  
11939.735 Million cell updates/sec

Title: US-10-005-907-1  
Perfect score: 3762  
Sequence: 1 gagaaacggactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004s: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3762	100.0	3762	6	ABN81319 Human mas
2	1068	28.4	1128	5	ABV20298 Human pro
3	1068	28.4	1128	5	ABV26130 Human pro
4	1068	28.4	1128	5	ABV26458 Human pro
5	1068	28.4	1128	5	ABV20617 Human pro
6	1068	28.4	1128	5	ABV20806 Human pro
7	1068	28.4	1128	5	ABV20439 Human pro
8	1068	28.4	1128	5	ABV26275 Human pro
9	1068	28.4	1128	5	ABV26653 Human pro
10	506	13.5	667	10	Adc30760 Human nov
11	394.2	10.5	433	5	ABV01270 Human pro
12	393.8	10.5	397	5	ABV10439 Human pro
13	392.2	10.4	421	5	ABV31276 Human pro
14	392.2	10.4	437	5	ABV31609 Human pro
15	392.2	10.4	437	5	ABV40245 Human pro
16	392.2	10.4	437	5	ABV40577 Human pro
17	391.8	10.4	404	5	ABV10104 Human pro
18	375	10.0	446	5	ABV00935 Human pro
19	342.8	9.1	359	5	ABV09593 Human pro
20	342.2	9.1	400	5	ABV00424 Human pro
21	342.2	9.1	401	5	ABV39737 Human pro

22	342.2	9.1	401	5	ABV30769	Human pro
23	342.2	9.1	401	5	ABV39958	Human pro
24	272.8	7.3	425	5	ABV09816	Human pro
25	244.2	6.5	296	5	ABV30990	Human pro
26	239.8	6.4	3337	4	AAK70219	Human imm
27	230.4	6.1	1982	5	ABV23873	Human pro
28	230.4	6.1	1982	5	ABV23839	Human pro
29	230.4	6.1	1982	5	ABV29715	Human pro
30	230.4	6.1	1982	5	ABV29752	Human pro
31	230.4	6.1	1982	5	ABV24256	Human pro
32	201.2	5.3	12278	5	ABA14585	Human ner
33	201.2	5.3	12278	5	ABA14664	Human ner
34	200.4	5.3	17239	4	AA36826	Human car
35	200.4	5.3	17239	10	ADE47520	Human car
36	200.4	5.3	17239	10	ABZ67711	Human sec
37	199.8	5.3	181257	12	ADF69677	Human SLC
38	199.4	5.3	2646	4	AA161012	Human pol
39	197.4	5.2	1743	4	AA162586	Human bre
40	197.4	5.2	1743	4	AA103368	Human rep
41	197.4	5.2	1746	4	AA162587	Human bre
42	197.4	5.2	1746	4	AA103369	Human rep
43	197.4	5.2	6685	4	AA307380	Human gen
44	197.4	5.2	6686	6	ABS73149	Human CLA
45	197	5.2	166043	12	ADL08127	Human gen

ALIGNMENTS

RESULT 1  
ABN81319

ID ABN81319 standard; CDNA; 3762 BP.

XX AC ABN81319;

XX DT 30-AUG-2002 (first entry)

XX DE Human mast cell related gene MC1 SEQ ID NO 1.

XX KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;

XX KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;

XX KW gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 25..432

XX FT /\*tag= a

XX FT /product= "MC1"

XX PN WO200246389-A2.

XX XX 13-JUN-2002.

XX PF 07-DEC-2001; 2001WO-US046180.

XX PR 08-DEC-2000; 2000US-0251835P.

XX PR 14-MAR-2001; 2001US-0275479P.

XX PR 28-MAR-2001; 2001US-0279115P.

XX PR 02-APR-2001; 2001US-0280143P.

XX XX (UNIO ) UCB SA.

XX XX Nocka K, Pirozzi G, Einstein R;

XX PI WPI; 2002-508560/54.

XX DR P-PSDB; ABB77569.

XX XX Novel isolated nucleic acids that are differentially expressed in mast

XX PT cells in patients with allergic hypersensitivity, encoding proteins

XX PT associated with mast cell regranulation and allergic hypersensitivity.

XX XX Claim 1; Page 95-97; 119pp; English.

**This Page Blank (uspto)**

QY 3408 AAGTTCCACAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGCAACC-AATGATCTGCT 3466  
Db 355 AAGTTCCACAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGCAACC-AATGATCTGCT 414  
QY 3467 TCGTATAATTAACAGTGTCTCTGAGTATTTGTAGCAATGTACCCCTTTCCATATTTATTTTG 3526  
Db 415 TCGTATAATTAACAGTGTCTCTGAGTATTTGTAGCAATGTACCCCTTTCCATATTTATTTTG 474  
QY 3527 TGTGTGTAAGGCTTCTTTTAGTCAATTAATAATTTTGGAGATTCATCTATGTTTAAATGTT 3586  
Db 475 GTGTGTGTAAGGCTTCTTTTAGTCAATTAATAAT-TGTGTGATTTCACTATGTTTAAATG-T 532  
QY 3587 CTATCAGTATGTTGATCACTTTACTTGTCTCAGCATATCAC-CATATAGATATACATATAAT 3645  
Db 533 CTATCAGTATGTTGATCACTTTACTTGTCTCAGCATATCACATATAGATTTACGTATAAT 592  
QY 3646 TTGTTAACTAATCACTGATGATATGTAGGATATTT--AAGTTTTCACATATTAAGAATA 3703  
Db 593 GTGTGTGTAAGGCTTCTTTTAGTCAATTAATAAT-TGTGTGATTTCACTATGTTTAAATG-T 532  
QY 3704 AAGTGGCTATAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3743  
Db 653 AAGTGGCTATAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 692

RESULT 15  
AI766257/c  
LOCUS wh68g04.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2385942 3',  
DEFINITION mRNA sequence.  
ACCESSION AI766257  
VERSION AI766257.1 GI:5232766  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 456)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Insert length: 1086 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 437.  
Location/Qualifiers  
1. .456  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2385942"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Kid11"  
/note="Organ: kidney; Vector: pTT73D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
plasmid DNA from the normalized library NCI CGAP Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.

FEATURES  
source

ORIGIN  
Query Match 11.9%; Score 446.6; DB 1; Length 456;  
Best Local Similarity 99.1%; Pred. No. 5.1e-53;  
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 3283 TCAAGTGTGATTATATACAGAAAAATGTAAACACCTGTGAAGGGTAGAGTTATAGAAATTTT 3342  
Db 456 TCAAGTGTGATTATATACAGAAAAATGTAAACACCTGTGAAGGGTAGAGTTATAGAAATTTT 397  
QY 3343 GTCRAATGTATTCACCCATGTAGTCACTCTTATTAAGAGACAGAACACGTACATCCTC 3402  
Db 396 GTCRAATGTATTCACCCATGTAGTCACTCTTATTAAGAGACAGAACACGTACATCCTC 337  
QY 3403 CCAGAAAGTTCACACAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGCAACC-AATGATC 3462  
Db 336 CCAGAAAGTTCACACAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGCAACC-AATGATC 277  
QY 3463 TGCCTTCGTATTAATTAATACTGTTCTAGATATTTGTAGCAATGTACCCCTTTCCATATTTAT 3522  
Db 276 TGCCTTCGTATTAATTAATACTGTTCTAGATATTTGTAGCAATGTACCCCTTTCCATATTTAT 217  
QY 3523 TTTGTGTGTGAAGGCTTCTTTTAGTCAATTAATAATTTTTCAGATTCACTATGTTTAA 3582  
Db 216 TTTGTGTGTGAAGGCTTCTTTTAGTCAATTAATAATTTTTCAGATTCACTATGTTTAA 157  
QY 3583 TGTCTTCATCAGTATGTTGTACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTAT 3642  
Db 156 TGTCTTCATCAGTATGTTGTACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTAT 97  
QY 3643 AATTTGTTAATCTAATCACTGATGGATATGTAGGATATTTAAAGTTTTCACATTTATGAAT 3702  
Db 96 AATTTGTTAATCTAATCACTGATGGATATGTAGGATATTTAAAGTTTTCACATTTATGAAT 37  
QY 3703 AAAGTGGCTATAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3735  
Db 36 AAAGTGGCTATAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 4

Search completed: November 18, 2004, 21:26:32  
Job time : 10884 secs

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
22709111  
12788976  
Contact: Volik SV  
Colin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7066  
Fax: 415 502 5665  
Email: svolik@cc.ucsf.edu  
This clone is available from Amplicon Express  
http://www.genomex.com  
Class: BAC ends.

FEATURES  
source

1. 757  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="MCF7\_1-3010"  
/sex="female"  
/clone\_lib="Human MCF7 breast cancer cell line library (MCF7\_1)"  
/note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 13.3%; Score 499.6; DB 8; Length 757;  
Best Local Similarity 98.2%; Pred. No. 1.6e-60;  
Matches 505; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3222 ATCTTTTATGCATATAATATTTCAACATTTTACATTTTATATTTTATATCTAGTTTATATCACTAGTTTAA 3281  
Db 1 ATCTTTTATGCATATAATATTTCAACATTTTACATTTTATATTTTATATCTAGTTTATATCACTAGTTTAA 60  
QY 3282 CTCAGTGTGATTATATACAGAAATGTACCACTGTAGGGTAGAGTTATAGATTT 3341  
Db 61 CTCAGTGTGATTATATACAGAAATGTACCACTGTAGGGTAGAGTTATAGATTT 120  
QY 3342 TGTCAAATGTATTCACCATGTAGTCCCTTTATGAAGACAGACACGTACATCCT 3401  
Db 121 TGTCAAATGTATTCACCATGTAGTCCCTTTATGAAGACAGACACGTACATCCT 180  
QY 3402 CCAGAAAGTTCACAGTGTCTTTTCCCTGTAGTTTCCAGTCTGGCAACCAATGAT 3461  
Db 181 CCAGAAAGTTCACAGTGTCTTTTCCCTGTAGTTTCCAGTCTGGCAACCAATGAT 240  
QY 3462 CTGCTTCGTATAATTAATTAATCTAGATATTTGTAGCAATGTACCTTTCCATATTTA 3521  
Db 241 CTGCTTCGTATAATTAATTAATCTAGATATTTGTAGCAATGTACCTTTCCATATTTA 300  
QY 3522 TTTTGTGTGTAAAGGTTCTTTTGTAGTCATTAATAATTTTGTAGATTCATCTATGTTA 3581  
Db 301 TTTTGTGTGTAAAGGTTCTTTTGTAGTCATTAATAATTTTGTAGATTCATCTATGTTA 360  
QY 3582 ATGTTTCTATCAGTAGTTGTATCTTACTGTCTCAGCATATCACCATATAGATATACTA 3641  
Db 361 ATGTTTCTATCAGTAGTTGTATCTTACTGTCTCAGCATATCACCATATAGATATACTA 420  
QY 3642 TAATTTGTTAACTAACTACGTAGGATATGTAGGATATTTAAGTATTTTGTACATTTAGAA 3701  
Db 421 TAATTTGTTAACTAACTACGTAGGATATGTAGGATATTTAAGTATTTTGTACATTTAGAA 480  
QY 3702 TAAAGTGGCTATAAATGAAAAAATAAAAAA 3735  
Db 481 TAAAGTGGCTATAAATGAAAAAATAAAAAA 514

RESULT 14  
BF243193  
LOCUS  
DEFINITION 601877334F1 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:4105665 5',  
EST 14-NOV-2000

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

mRNA sequence.  
BF243193  
BF243193.1 GI:11157122  
EST.  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 848)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM984 row: c column: 10  
High quality sequence stop: 546.  
Location/Qualifiers  
1. 848  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4105665"  
/tissue\_type="from acute myelogenous leukemia"  
/lab\_host="DHI0B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 55"  
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);  
Site 1: SfiI (ggcgcctggcc); Site 2: SfiI  
(ggcgtatggcc); Double-stranded cDNA was prepared from  
cell line RNA. 5' and 3' adaptors were used in cloning as  
follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and  
3' adaptor sequence:  
5'-ATTCTAGAGCGCGCCGACATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size  
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained  
inserts by PCR. This library was enriched for full-length  
clones and was constructed by Clontech Laboratories (Palo  
Alto, CA)."

FEATURES  
source

ORIGIN

Query Match 13.0%; Score 487.6; DB 2; Length 848;  
Best Local Similarity 90.4%; Pred. No. 7.6e-59;  
Matches 633; Conservative 0; Mismatches 54; Indels 13; Gaps 10;

QY 3049 CATCAATGCTTCTCGTGGTTTACCATAAGCCACAGACAGATGCTTTAATCTTTCCGAGAT 3108  
Db 1 CATCAATGCTTCTCGTGGTTTACCATAAGCCACAGACAGATGCTTTAATCTTTCCGAGAT 60  
QY 3109 CTAGTTTTTCAGAAAGCAGGATTTAAGAAATGTAATCTATCTGTTATGAAGAAC 3168  
Db 61 CTAGTTTTTCAGAAAGCAGGATTTAAGAAATGTAATCTATCTGTTATGAAGAAC 120  
QY 3169 AATAGAAATCATTTGCTGTAAGTCTTTTAACTCTTAATCTATCTGTTATGAAGAAC 3168  
Db 61 CTAGTTTTTCAGAAAGCAGGATTTAAGAAATGTAATCTATCTGTTATGAAGAAC 120  
QY 3169 AATAGAAATCATTTGCTGTAAGTCTTTTAACTCTTAATCTATCTGTTATGAAGAAC 3168  
Db 121 AATAGAAATCATTTGCTGTAAGTCTTTTAACTCTTAATCTATCTGTTATGAAGAAC 3228  
QY 3229 ATGCATATAAATTTTGAACATTTTACATTTTATATTTTAACTCAGTTTCTTACTCAAGT 3288  
Db 179 ATGCATATAAATTTTGAACATTTTACATTTTATATTTTAACTCAGTTTCTTACTCAAGT 236  
QY 3289 GTGATTATATCAAGAAATGTAACCTGTAAAGGTAGAGTTATGAAGAATTTTGTCAA 3348  
Db 237 GTGATTATATCAAGAAATGTAACCTGTAAAGGTAGAGTTATGAAGAATTTTGTCAA 294  
QY 3349 TGTATTCCCAATGTAGTACCTCTTATGAGAGACAGAA-CAGGTACATCTCTCCAGA 3407  
Db 295 TGTATTCCCAATGTAGTACCTCTTATGAGAGACAGAA-CAGGTACATCTCTCCAGA 354

```
Db 120 TTAGACAACTACTGTACAAATGAAATAAATTTCTTAATCCCTGACTAACTGAATGAC 179
Qy 761 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAG 820
Db 180 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAG 239
Qy 821 GAAGGAGGTGAGACACACCTGTTATACCCCTTCCCTTTGGAGTTTATGACAAAGTGA 880
Db 240 GAAGGAGGTGAGACACACCTGTTATACCCCTTCCCTTTGGAGTTTATGACAAAGTGA 298
Qy 881 CCAGGATGAGTGCATGAAGACTGTGAATAGACTGATTTGGCAATGAAGAGTCCCAATTC 940
Db 299 CCAGGATGAGTGCATGAAGACTGTGAATAGACTGATTTGGCAATGAAGAGTCCCAATTC 358
Qy 941 AACCTGACTCTGGTGTAGATCAACACTGTCTGAGGAGTTCATCTATGAGACTTTGCT 1000
Db 359 AACCTGACTCTGGTGTAGATCAACACTGTCTGAGGAGTTCATCTATGAGACTTTGCT 418
Qy 1001 ACATAACAGAGACCTGGTTTCCACAAACCCCTTTATTTAGCTTAAAGCAATCTTTTCTAC 1060
Db 419 ACATAACAGAGACCTGGTTTCCACAAACCCCTTTATTTAGCTTAAAGCAATCTTTTCTAC 478
Qy 1061 TGACTTCTTAAGTCTTTAGACA--AAGCTTAACTCTTTCAACCAAT--GCCAATCAGACA 1117
Db 479 TGACTTCTTAAGTCTTTAGACAAGAGCTTTAACTCTTTCAACCAATTTGCCAATCAGACA 538
Qy 1118 AACTTTGAA--TCTACCTATGACC--TGTAAGCTCTCTCTGCTTGAAGATCTTGCCTC-- 1172
Db 539 AACTTTGAA--TCTACCTATGACC--TGTAAGCTCTCTCTGCTTGAAGATCTTGCCTC-- 598
Qy 1173 -TTTAAGCTGAA--CCGATGTGACATTT--CCATTTAATGATTTTATGCTTTGTTGTA 1226
Db 599 TTTAAGCTGAA--CCGATGTGACATTT--CCATTTAATGATTTTATGCTTTGTTGTA 658
Qy 1227 A--CTCCTGTCTCCCTAAATGATATAAAGTAAAC--GGTGACCTGACCACTCAGGCACA 1283
Db 659 AACTCTCTGCTCTGCTAAATGTTTAAAGTCAACGGGTGACCTGACCACTCAGGCACA 718
Qy 1284 CTTTCTCA--GGACCTCTGAGAGTGTATCCGAGGC 1317
Db 719 CTTTCTCAAGAACTCTCTGGAAGTGGATCCGAGGC 754

RESULT 12
AW65849/C
LOCUS
DEFINITION
hi94f08.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:2979975 3', mRNA sequence.
EST 06-APR-2000
ACCESSION
AW65849
VERSION
AW65849.1 GI:7458398
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 504)
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 451.
FEATURES
source
1..504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2979975"
/lab_host="DH10B"
```

/clone lib="Soares NFL T GBC S1"  
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBLH19W, testis NHT, and B-cell  
NCI CGAP GCB1) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
1.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo. "

ORIGIN

```
Query Match 13.3%; Score 500.8; DB 2; Length 504;  
Best Local Similarity 99.6%; Pred. No. 1.3e-60;  
Matches 502; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 3219 CTTATCTTTTATGCATATAAATATTTGAACATTTTACATTTGTTTATTTTAAATCAGTT 3278  
Db 504 CTTATCTTTTATGCATATAAATATTTGAACATTTTACATTTGTTTATTTTAAATCAGTT 445  
Qy 3279 TTACTCAAGTGTGATATATACAGAAATGTAACACCTGTAAGGTAGAGTATAGAA 3338  
Db 444 TTACTCAAGTGTGATATATACAGAAATGTAACACCTGTAAGGTAGAGTATAGAA 385  
Qy 3339 TTTTGTCAATGTATTTTACCCATGTAGTCACCTCTTATGAAGAGACAGAACAGTACAT 3398  
Db 384 TTTTGTCAATGTATTTTACCCATGTAGTCACCTCTTATGAAGAGACAGAACAGTACAT 325  
Qy 3399 CTTCCAGAAAGTTCCACAGTGTCTCTTTTCCCTGAGTTTCCACAGTCTGGCAACCAAT 3458  
Db 324 CTTCCAGAAAGTTCCACAGTGTCTCTTTTCCCTGAGTTTCCACAGTCTGGCAACCAAT 265  
Qy 3459 GATCTCTCTGATATAATATACTGTTCTAGATATTTTGTAGCAATGTACCTTTCCATAT 3518  
Db 264 GATCTCTCTGATATAATATACTGTTCTAGATATTTTGTAGCAATGTACCTTTCCATAT 205  
Qy 3519 TTATTTTGTGTGTGAAGCTTCTTTTGTAGTCAATATAATTTTGTAGATTTTGTATGT 3578  
Db 204 TTATTTTGTGTGTGAAGCTTCTTTTGTAGTCAATATAATTTTGTAGATTTTGTATGT 145  
Qy 3579 TTAATGTTCTATCAGTAGTTGTACATCTTTTCTCAGCATATCACCATATAGATATA 3638  
Db 144 TTAATGTTCTATCAGTAGTTGTACATCTTTTCTCAGCATATCACCATATAGATATA 85  
Qy 3639 CTATAATTTGTTAACTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 3698  
Db 84 CTATAATTTGTTAACTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 25  
Qy 3699 GAATAAGTGGCTATAAATGAAA 3722  
Db 24 GAATAAGTGGCAATAAATGAAA 1
```

RESULT 13  
BZ600841  
LOCUS  
DEFINITION  
WHAAJ89TR Human MCF7 breast cancer cell line library (MCF7\_1) Homo  
sapiens genomic clone MCF7\_1-3010, genomic survey sequence.  
ACCESSION  
BZ600841  
VERSION  
BZ600841.1 GI:31509303  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
REFERENCE  
1 (bases 1 to 757)  
AUTHORS  
Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q.,  
Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,  
Gray, J.W. and Collins, C.  
TITLE  
End-sequence profiling: Sequence-based analysis of aberrant genomes



```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM993 row: a column: 01
High quality sequence stop: 518.
FEATURES
source
1..527
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfII (ggcgctcgcc); Site 2: SfII
(ggcattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CAGGCGCATTTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

```

Query Match 13.5%; Score 506; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.3e-61;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACTGTGAAAGATCGGAATTTATCTCTCGCAAACTCAGTTGCTG 60
DB 21 GAGAAACCGAGTCACTGTGAAAGATCGGAATTTATCTCTCGCAAACTCAGTTGCTG 80
QY 61 GGAGAGAAATCAAAAGAACCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 120
DB 81 GGAGAGAAATCAAAAGAAAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 140
QY 121 ATGACTACATTTCAAGAAAGAAAGTCAAGATCAAGATCAAGAAAGAAAGTTCATCC 180
DB 141 ATGACTACATTTCAAGAAAGAAAGTCAAGATCAAGATCAAGAAAGAAAGTTCATCC 200
QY 181 ACTTCTAATCAGGAAACGAGATGCGAGTGTCTGGAAGTGTCTACATGTCATT 240
DB 201 ACTTCTAATCAGGAAACGAGATGCGAGTGTCTGGAAGTGTCTACATGTCATT 260
QY 241 AATCATATCCCCATCAGAGATCCTCCCTGAGCTCCAATGATGATGGCTATGAGACATT 300
DB 261 AATCATATCCCCATCAGAGATCCTCCCTGAGCTCCAATGATGATGGCTATGAGACATT 320
QY 301 GACTCCCTCACAAGAAAGTGAGACAGTTTAGAGAAAGGTTCAGACAGAAATATGCCCTT 360
DB 321 GACTCCCTCACAAGAAAGTGAGACAGTTTAGAGAAAGGTTCAGACAGAAATATGCCCTT 380
QY 361 CTTAGGACTTCTGTAGTAGGCTTCTCTCTGACCCCATGAGCATGATTAAGAGTTGTG 420
DB 381 CTTAGGACTTCTGTAGTAGGCTTCTCTCTGACCCCATGAGCATGATTAAGAGTTGTG 440
```

```

QY 421 TTTCCACACTAAATTCCTAAGCTGCTTTATCACCCTCCAGCAATGAAGACAATGCAGAA 480
DB 441 TTTCCACACTAAATTCCTAAGCTGCTTTATCACCCTCCAGCAATGAAGACAATGCAGAA 500
QY 481 TAGCAGACTCTGGGGAAGTGTTCAC 506
DB 501 TAGCAGACTCTGGGGAAGTGTTCAC 526

RESULT 11
BF240893 809 bp mRNA linear EST 14-NOV-2000
LOCUS 60187544F2 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4091963 5',
DEFINITION mRNA sequence.
ACCESSION BF240893
VERSION BF240893.1 GI:11154817
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 809)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM948 row: h column: 12
High quality sequence stop: 651.
FEATURES
source
1..809
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4091963"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfII (ggcgctcgcc); Site 2: SfII
(ggcattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CAGGCGCATTTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

```

Query Match 13.3%; Score 502; DB 2; Length 809;
Best Local Similarity 90.5%; Pred. No. 7.4e-61;
Matches 684; Conservative 0; Mismatches 50; Indels 22; Gaps 13;

QY 582 CTGATATGTGGCATCTCTGTGGCTTAGGTGAATCATAGAAATTCACAAATGACCTAAA 641
DB 1 CTGATATGTGGCATCTCTGTGGCTTAGGTGAATCATAGAAATTCACAAATGACCTAAA 60
QY 642 ATATTCTATGTGTTTTGCTTGAAGTTTGAAGCATGGAGGTGATAAAAAACTTTC 701
DB 61 ATATTCTATGTG-TTTTGCTTGAAGTTTGAAGCATGGAGGTGATAAAAAACTTTC 119
QY 702 TTAGGACAAATATGA-BAATGAAATAAATTTCTTAATCCCTGACTGAATGATGAC 760
```



```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 728)
NTH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM210 row: n column: 16
High quality sequence stop: 604.
FEATURES
    source
        1..728
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:30414375"
            /tissue_type="Pooled"
            /lab_host="DH10B (TI phage-resistant)"
            /clone_lib="NIH_MGC_191"
            /notes="Vector: pDNR-LIB; Site 1: Sfil (ggccattatggcc);
            Site 2: Sfil (ggccgctggcc); Library is oligo-dT primed
            and directionally cloned. PBMC - Peripheral Blood
            Mononuclear Cells. RNA was pooled from 3/6hour stimulation
            with PMA adn ionomycin. 5' and 3' adaptors were used in
            cloning as follows: 5' adaptor sequence:
            5'-CAGCGCCATATGCCC-3' and 3' adaptor sequence:
            5'-ATTCTAGAGCCGAGCGCGGACATG-dt(30)BN-3' (where B = A,
            C, or G and N = A, C, G, or T). Average insert size 1.69
            kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
            by PCR. This library was enriched for full-length clones
            and was constructed by Clontech Laboratories (Palo Alto,
            CA). Note: this is a NIH_MGC Library."
ORIGIN
    Query Match 16.3%; Score 612.2; DB 6; Length 728;
    Best Local Similarity 98.4%; Pred. No. 2.8e-76;
    Matches 617; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1835 ATCTTACCGAAGGTGGATGTGTAGTTCTTCCCAAGAGTTTTTAAATGATATGCCAGCTTCCTAAT 1954
Db 3 ATCTTACCGAAGGTGGATGTGTAGTTCTTCCCAAGAGTTTTTAAATGATATGCCAGCTTCCTAAT 1954
QY 1895 ATCCATCTAATATGCTTTCTTTCCCAAGAGTTTTTAAATGATATGCCAGCTTCCTAAT 1954
Db 63 ATCCATCTAATATGCTTTCTTTCCCAAGAGTTTTTAAATGATATGCCAGCTTCCTAAT 122
QY 1955 TTGGAGACAAAAGCCCTTAATTTGACATGCAATGCAATCATATATATTTTGTATAGTTACAG 2014
Db 123 TTGGAGACAAAAGCCCTTAATTTGACATGCAATGCAATCATATATATTTTGTATAGTTACAG 182
QY 2015 TATACAGATTGAGTATCCCTTATGATGAGATGCTTTGGAGCAGAGGTTTGGATTTTCAG 2074
Db 183 TATACAGATTGAGTATCCCTTATGATGAGATGCTTTGGAGCAGAGGTTTGGATTTTCAG 242
QY 2075 ATTTATTTTGGATTTTGGAAATATTTCCATACATATATATGAGAGAGTTGGAAAATGGGAT 2134
Db 243 ATTTATTTTGGATTTTGGAAATATTTCCATACATATATATGAGAGAGTTGGAAAATGGGAT 302
QY 2135 TCAAGTCTAATCATAAAATTCACCTATGTTTGTATATACACCTTATCTGAATAGCCTGAAG 2194
Db 303 TCAAGTCTAATCATAAAATTCACCTATGTTTGTATATACACCTTATCTGAATAGCCTGAAG 362

```

---

```

QY 2195 GTAATTTTATACAAATATTTTAAATAATTTTATGCTGAAAACAGAGTTTGGCACAATTGGA 2254
Db 363 GTAATTTTATACAAATATTTTAAATAATTTTATGCTGAAAACAGAGTTTGGCACAATTGGA 422
QY 2255 CCATCAGAAAGCAGAGTGTCACTATTTCAAGTCAGTGCCTCAAAAAGTTTTCAGATGTAA 2314
Db 423 CCATCAGAAAGCAGAGTGTCACTATTTCAAGTCAGTGCCTCAAAAAGTTTTCAGATGTAA 482
QY 2315 GCTGGTGATGAGTTTCATGCCAGTATCCGAGTACTTTTGGGAAAGCCCAAGACAGGTGGATC 2374
Db 483 GCTGGTGATGAGTTTCATGCCAGTATCCGAGTACTTTTGGGAAAGCCCAAGACAGGTGGATC 542
QY 2375 TCTTGAGCCAGAGTTTGGAGCCAGTCTGACCAACACACAGTGCAGCTCTGTTCTACAAA 2434
Db 543 TCTTGAGCCAGAGTTTGGAGCCAGTCTGACCAACACACAGTGCAGCTCTGTTCTACAAA 602
QY 2435 TAATTTAAAAAATTAGCCAGGTGGTG 2461
Db 603 TATTAATAAATTNAGCAGTGTGTGTG 629

```

---

```

RESULT 8
AGI82348
LOCUS Pan troglodytes DNA, clone: RP43-055101.TJ, genomic survey
DEFINITION
AGI82348
VERSION AGI82348.1 GI:16712028
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
    FUJIIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
    TOTOKI, Y., WATANABE, H. and SAKAKI, Y.
    BAC end sequences of Library RPCI-43
    Unpublished
    2 (bases 1 to 643)
    FUJIIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
    TOTOKI, Y., WATANABE, H. and SAKAKI, Y.
    Direct Submission
    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
    and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
    1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
    (E-mail: chimpanse@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
    Tel: 81-45-503-9111, Fax: 81-45-503-9170)
    Clones are derived from the chimpanzee BAC library RPCI-43. This BAC
    end was generated during the R&D process and may have higher chance
    of clone tracking errors.
    PRIMERS
    Sequencing: TJ
    LIBRARY
        Vector : pBACe3.6
        R.Site 1 : EcoRI
        R.Site 2 : EcoRI.
        Location/Qualifiers
            1..643
                /organism="Pan troglodytes"
                /mol_type="genomic DNA"
                /db_xref="taxon:9598"
                /clone="RP43-055101.TJ"
                /sex="male"
                /cell_type="lymphocytes"
                /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN
    Query Match 15.2%; Score 570.4; DB 9; Length 643;
    Best Local Similarity 95.4%; Pred. No. 2.1e-70;
    Matches 599; Conservative 0; Mismatches 26; Indels 3; Gaps 1;
QY 2751 GAATTTCTATGCGCTTCTGTGTGTTTATGTGCGGAAAGGAAACAAATGAAATTTTG 2810

```

cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTAGAGCGGCGGACATG-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

QY 821 GAAGGAGCTGAGACACACCTTGTATATACCCCTCCCTTTGGAGTTTATGCACAAGTGA 880  
Db 303 GAAGGAGCTGAGACACACCTTGTATATACCCCTCCCTTTGGAGTTTATGCACAAGTGA 362  
QY 881 CCAGGATGAGTCATAAGACTGATGAATAGACTGATTTGGCAATAAGAGTCCCAATTC 940  
Db 363 CCAGGATGAGTCATAAGACTGATGAATAGACTGATTTGGCAATAAGAGTCCCAATTC 422  
QY 941 AACCTGACTCTGCTGATGATACACACTGCTGAGGAGTTCATCTATGAGACTTTGCT 1000  
Db 423 AACCTGACTCTGCTGATGATACACACTGCTGAGGAGTTCATCTATGAGACTTTGCT 482  
QY 1001 ACATTAACAGACCTTGTTCACACACCCCTTTATTTAGCTTAAGCATCTTTCTAC 1060  
Db 483 ACATTAACAGACCTTGTTCACACACCCCTTTATTTAGCTTAAGCATCTTTCTAC 542  
QY 1061 TGACTTCTTAAGTCTTTAGACAAAGCTTAACCTTTCAACCAATTCGCCAATCAGACAAAC 1120  
Db 543 TGACTTCTTAAGTCTTTAGACAAAGCTTAACCTTTCAACCAATTCGCCAATCAGACAAAC 602  
QY 1121 TTTGAATCTACTATGACTGATGAGCTCTCTCTGCTTCAGATCTTGCCCTCTTTAAGCT 1180  
Db 603 TTTGAATCTACTATGACTGATGAGCTCTCTCTGCTTCAGATCTTGCCCTCTTTAAGCT 662  
QY 1181 GAACC--GATGTGACCTTCCATTTAATGATTAT-GTCTTTGCTTGTA--CTCCTGT 1234  
Db 663 GAACCCGATGTCCTTTCCCATTTAATGATTATGAGCTTTGCTTGTAACCTCTCTGTC 722  
QY 1235 CTCCTTAAATGTATAAA 1253  
Db 723 TCCCTTAAATGTATAAA 741

RESULT 6  
LOCUS BUI55429 786 bp mRNA linear EST 03-SBP-2002  
DEFINITION AGENCOURT\_7979368 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:6108338  
5', mRNA sequence.  
ACCESSION BUI55429  
VERSION BUI55429.1 GI:22668961  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA library preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Cloning distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM2352 row: h column: 03  
High quality sequence stop: 399.  
Location/Qualifiers  
1..786  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6108338"  
/tissue="from acute myelogenous leukemia"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_55"  
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: Sfil (ggccgctcgcc); Site 2: Sfil (ggcattatggc); Double-stranded cDNA was prepared from

FEATURES  
source  
1..786  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6108338"  
/tissue="from acute myelogenous leukemia"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_55"  
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: Sfil (ggccgctcgcc); Site 2: Sfil (ggcattatggc); Double-stranded cDNA was prepared from

RESULT 7  
LOCUS CD639433  
DEFINITION AGENCOURT\_14532881 NIH\_MGC\_191 Homo sapiens cDNA clone  
IMAGE:30414375 5', mRNA sequence.  
ACCESSION CD639433  
VERSION CD639433.1 GI:31806240  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

Query Match 16.7%; Score 628.6; DB 5; Length 786;  
Best Local Similarity 97.7%; Pred. No. 1.4e-78;  
Matches 648; Conservative 0; Mismatches 14; Indels 1; Gaps 1;  
ORIGIN  
QY 2609 GTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTTAGGAGTTTCAACCTGACTGACCT 2668  
Db 3 GTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTTAGGAGTTTCAACCTGACTGACCT 62  
QY 2669 TTTAGTCTTCAAGCAAGCAATTAATCAATAGGTGGAGCTCCAGATAAATCAATTTCTGTATAT 2728  
Db 63 TTTAGTCTTCAAGCAAGCAATTAATCAATAGGTGGAGCTCCAGATAAATCAATTTCTGTATAT 122  
QY 2729 ACATTTGCTCTCTATTCAAGCAATTTCTATTGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 2788  
Db 123 ACATTTGCTCTCTATTCAAGCAATTTCTATTGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 182  
QY 2789 AAGGGAACCAATAGAAAATTTTGCATTTCTAGAAAAGTCAATCTGTCAAAATATGTCAGT 2848  
Db 183 AAGGGAACCAATAGAAAATTTTGCATTTCTAGAAAAGTCAATCTGTCAAAATATGTCAGT 242  
QY 2849 CTGTAGATATTAGCCAAATTTTAGGAAAATGACAAAATTTTACCTTTCTGCTGCTGCTGCTG 2908  
Db 243 CTGTAGATATTAGCCAAATTTTAGGAAAATGACAAAATTTTACCTTTCTGCTGCTGCTGCTG 302  
QY 2909 TAGCTGTTTATGATATATAAATACCTTATTGTAATAAATAAATTTTAAATTTTGAAGTAA 2968  
Db 303 TAGCTGTTTATGATATATAAATACCTTATTGTAATAAATAAATTTTAAATTTTGAAGTAA 362  
QY 2969 ATCTGGAATTTATCAGAGAGGCGGCAAGCAATAGGTTTAATAACAGATTGATTGTTAGAA 3028  
Db 363 ATCTGGAATTTATCAGAGAGGCGGCAAGCAATAGGTTTAATAACAGATTGATTGTTAGAA 422  
QY 3029 GGAACCTTGAAATCCAAAGAGCATCAATGCTCTTCTGCTGCTTCCACCAAGCCACACAGAGA 3088  
Db 423 GGAACCTTGAAATCCAAAGAGCATCAATGCTCTTCTGCTGCTTCCACCAAGCCACACAGAGA 482  
QY 3089 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGAGGAGTTTGAAGAAATGTAATAT 3148  
Db 483 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGAGGAGTTTGAAGAAATGTAATAT 542  
QY 3149 CTTATGCTGTTATGAGAACCAATAGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3207  
Db 543 CTTATGCTGTTATGAGAACCAATAGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602  
QY 3208 ATTTTGTGAAGCTTATCTTTTATGATATAAATAAATTTTGAACATTTTACATTTGTTATAT 3267  
Db 603 ATTTTGTGAAGCTTATCTTTTCTGCTATAAATAAATTTTGAACATTTTACATTTGTTATAT 662  
QY 3268 TTT 3270  
Db 663 TTT 665

/lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_55"  
 /note="Organ: Bone marrow; Vector: pDNR-LIB (Clontech);  
 Site 1: SfiI (ggcgcctcgcc); Site 2: SfiI  
 (ggccattatggcc); Double-stranded cDNA was prepared from  
 cell line RNA. 5' and 3' adaptors were used in cloning as  
 follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and  
 3' adaptor sequence:  
 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo  
 Alto, CA)."

## ORIGIN

Query Match 18.4%; Score 692.4; DB 5; Length 785;  
 Best Local Similarity 97.0%; Pred. No. 1.6e-87;  
 Matches 738; Conservative 0; Mismatches 17; Indels 6; Gaps 3;

QY 2055 AGAAGTGTGGATTTTCAGATTTATTTTGGATTTGGAAATTTCCATACATATAATG 2114  
 Db |||||||  
 3 AGAAGTGTGGATTTTCAGATTTATTTTGGATTTGGAAATTTCCATACATATAATG 62

QY 2115 AGAGAGTTGGAATGGGATTCAGTCTAATCATATAAAATTCACATTATGTTGATATAC 2174  
 Db |||||||  
 63 AGAGAGTTGGAATGGGATTCAGTCTAATCATATAAAATTCACATTATGTTGATATAC 122

QY 2175 CTTATCTGAATAGCTGAAGTAATTTTATACAAATTTTAAATTAATTTTATGCTGAAA 2234  
 Db |||||||  
 123 CTTATCTGAATAGCTGAAGTAATTTTATACAAATTTTAAATTAATTTTATGCTGAAA 182

QY 2235 CAGAGTTTGGCAGATTTGGCCATCAGAACAGAGTGTCACTATTTCAGTCAAGTGCT 2294  
 Db |||||||  
 183 CAGAGTTTGGCAGATTTGGCCATCAGAACAGAGTGTCACTATTTCAGTCAAGTGCT 242

QY 2295 CAAAAGTTTCAGATGTTAAGCTGGTGTATGAGTTCATGCCAGTGATCCGAGTACTTTGG 2354  
 Db |||||||  
 243 CAAAAGTTTCAGATGTTAAGCTGGTGTATGAGTTCATGCCAGTGATCCGAGTACTTTGG 302

QY 2355 GAAGCCAGAGCAGTGATCTTTGAGCCAGGAGTTTGAGCCAGAGCTGCACAAACAG 2414  
 Db |||||||  
 303 GAAGCCAGAGCAGTGATCTTTGAGCCAGGAGTTTGAGCCAGAGCTGCACAAACAG 362

QY 2415 TGAGACCTGTTTCTACAAATTAATTAATAATTAAGCAGTGTGCTGTGTCACACCTGTA 2474  
 Db |||||||  
 363 TGAGACCTGTTTCTACAAATTAATTAATAATTAAGCAGTGTGCTGTGTCACACCTGTA 422

QY 2475 GTCCAGGTACTCAGAGGCTGAGGTAGTAGTGGATTGTTTGAGACTGGGAGGTTGAGGCTG 2534  
 Db |||||||  
 423 GTCCAGGTACTCAGAGGCTGAGGTAGTAGTGGATTGTTTGAGACTGGGAGGTTGAGGCTG 482

QY 2535 AACTGAGCCAGGATCTTGCCACACATTCAGCTTGGGCAACAGAGTGAGACCTGTCTC 2594  
 Db |||||||  
 483 AACTGAGCCAGGATCTTGCCACACATTCAGCTTGGGCAACAGAGTGAGACCTGTCTC 542

QY 2595 AAAAAAAGTTTCAGATTTTGGAGCATTTCAGATCTTCAGATTAGGATTTTCA 2654  
 Db |||||||  
 543 AAAAAAAGTTTCAGATTTTGGAGCATTTCAGATCTTCAGATTAGGATTTTCA 602

QY 2655 ACTGTACTGACCTTTTGTGTCATTCAGACATTAATCAATAGTGGACTCCAGATAACT 2714  
 Db |||||||  
 603 ACTGTACTGACCTTTTGTGTCATTCAGACATTAATCAATAGTGGACTCCAGATAACT 662

QY 2715 CATTTGCTGTATACACA-TTTTGGCTCTCTATTCAACGAATCTT--ATGCCCTCTTTGG 2771  
 Db |||||||  
 663 CATTTGCTGTATACACA-TTTTGGCTCTCTCTATTCAACGAATCTT--ATGCCCTCTTTGG 722

QY 2772 GTGATTTTAA---TGTCGGAAGGGAACCAATAGAAATTTT 2809  
 Db |||||||  
 723 GTGATTTTAA---TGTCGGAAGGGAACCAATAGAAATTTT 763

RESULT 5  
 BUI55309  
 LOCUS  
 DEFINITION  
 AGENCOURT 7951175 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:6107719  
 5', mRNA sequence.  
 ACCESSION  
 BUI55309  
 VERSION  
 BUI55309.1 GI:22668841  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (Bases 1 to 741)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: rs98@nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM2350 row: n column: 08  
 High quality sequence stop: 512.

FEATURES  
 source

1..741  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6107719"  
 /tissue\_type="from acute myelogenous leukemia"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_55"  
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);  
 Site 1: SfiI (ggcgcctcgcc); Site 2: SfiI  
 (ggccattatggcc); Double-stranded cDNA was prepared from  
 cell line RNA. 5' and 3' adaptors were used in cloning as  
 follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and  
 3' adaptor sequence:  
 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo  
 Alto, CA)."

## ORIGIN

Query Match 17.9%; Score 673; DB 5; Length 741;  
 Best Local Similarity 97.2%; Pred. No. 8.4e-85;  
 Matches 718; Conservative 0; Mismatches 15; Indels 6; Gaps 3;

QY 521 GAAACATCTCTTCTGGCTAAAGTTTAGAAAATTAATCTTATATATATCTTATAGCAAC 580  
 Db |||||||  
 3 GAAACATCTCTTCTGGCTAAAGTTTAGAAAATTAATCTTATATATATCTTATAGCAAC 62

QY 581 TCTGATATCTGTCATCTCTGCTTAGTGAATCATAGAAATTCACAAATGACCAATGACCTAA 640  
 Db |||||||  
 63 TCTGATATCTGTCATCTCTGCTTAGTGAATCATAGAAATTCACAAATGACCAATGACCTAA 122

QY 641 AATATCTATGTGTTTTTCTTGTAAAGTTTTCAGGACATGAGGAGTGAATAAAAAAATTTT 700  
 Db |||||||  
 123 AATATCTATGTGTTTTTCTTGTAAAGTTTTCAGGACATGAGGAGTGAATAAAAAAATTTT 182

QY 701 CTTAGACAAATATGTAAATCAAAATTAATTTCTAATCCCTGACTAATCTGAATGAGAC 760  
 Db |||||||  
 183 CTTAGACAAATATGTAAATCAAAATTAATTTCTAATCCCTGACTAATCTGAATGAGAC 242

QY 761 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAGACTGAATTCCTGGCCATGATAG 820  
 Db |||||||  
 243 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAGACTGAATTCCTGGCCATGATAG 302





Query Match 18.5%; Score 696.8; DB 5; Length 749;  
 Best Local Similarity 96.3%; Pred. No. 3.9e-88;  
 Matches 704; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 3029 GGAACGTTGAAATCCAGAGCATCAATGCTCTTCTGGTGTTCACCATAGCCACAGCAGA 3088  
 DB 3 GGAACGTTGAAATCCAGAGCATCAATGCTCTTCTGGTGTTCACCATAGCCACAGCAGA 62

QY 3089 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGCAGGATTTAAGAAATGTAACCTAT 3148  
 DB 63 TGCCTTAATCTTCCGAGATCTAGTTTTCAGCAAGCAGGATTTAAGAAATGTAACCTAT 122

QY 3149 CTTATGCTGTATGAAGCAACATGAATCAATGCTGTATAGTGCTTTTAACTGTAA 3208  
 DB 123 CTTATGCTGTATGAAGCAACATGAATCAATGCTGTATAGTGCTTTTAACTGTAA 182

QY 3209 TTTTGTGAAGCTATCTTTTATGCAATAAATAATTTGAACATTTTACATGTTATATTT 3268  
 DB 183 TTTTGTGAAGCTATCTTTTATGCAATAAATAATTTGAACATTTTACATGTTATATTT 242

QY 3269 TTAATCAGTTTACTCAAGTGTGATTTATACAGAAAATGTAACCATGTAAGGCTAGA 3328  
 DB 243 TTAATCAGTTTACTCAAGTGTGATTTATACAGAAAATGTAACCATGTAAGGCTAGA 302

QY 3329 GTTATAAGAAATTTTGTCAAATGTATTCACCCATGTAGTCACCTCCTTATGAAGAGCAGA 3388  
 DB 303 GTTATAAGAAATTTTGTCAAATGTATTCACCCATGTAGTCACCTCCTTATGAAGAGCAGA 362

QY 3389 ACAGTACATCTCCAGAAAGTTCCACAGTGCTCTTTTCCCTGAGTTTCCACAGTCCT 3448  
 DB 363 ACAGTACATCTCCAGAAAGTTCCACAGTGCTCTTTTCCCTGAGTTTCCACAGTCCT 422

QY 3449 GGCAACCAATGATCTGCTGTATATTAATTAACCTGTAGATATTTGTAGCAATGACC 3508  
 DB 423 GGCAACCAATGATCTGCTGTATATTAATTAACCTGTAGATATTTGTAGCAATGACC 482

QY 3509 CTTTCCATATTTATTTGTGTGTGAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGAT 3568  
 DB 483 CTTTCCATATTTATTTGTGTGTGAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGAT 542

QY 3569 TCATCTATGTTTAAATGTTTATCAGTAGTGTGACATCTTACTTGTCTCAGCATATCACA 3628  
 DB 543 TCATCTATGTTTAAATGTTTATCAGTAGTGTGACATCTTACTTGTCTCAGCATATCACA 602

QY 3629 TATAGATATATAATTTGTTTAACTTAACCTGATGATATGTAGATATTTAAATTT 3688  
 DB 603 TATAGATATATAATTTGTTTAACTTAACCTGATGATATGTAGATATTTAAATTT 662

QY 3689 TTGACATTTATGAATGAAGTGGCTATAAATGAAAGAAAAAAGAAAAAAGAAAAA 3748  
 DB 663 TTGACATTTATGAATGAAGTGGCTATAAATGAAAGAAAAAAGAAAAAAGAAAAA 722

QY 3749 AAAAAAANA 3759  
 DB 723 NNNNNANANA 733

RESULT 2  
 BO441475  
 LOCUS  
 DEFINITION BO441475 782 bp mRNA linear EST 24-MAY-2002  
 AGENCOURT\_7838612 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6101435  
 5', mRNA sequence.  
 ACCESSION BO441475  
 VERSION BO441475.1 GI:21180551  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 782)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILML at:  
<http://image.llnl.gov>  
 Plate: LLCM2334 row: h column: 12  
 High quality sequence stop: 561.

## FEATURES

Location/Qualifiers  
 1..782  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6101435"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /clone\_lib="NIH\_MGC\_82"  
 /notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgctcgcc); Site 2: SfiI (ggccattatgccc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAGCGCATATGCGC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGCGCGGCGGAGCATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo  
 Alto, CA)."

## ORIGIN

Query Match 18.5%; Score 695.8; DB 5; Length 782;  
 Best Local Similarity 97.0%; Pred. No. 5.3e-88;  
 Matches 709; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 3029 GGAACGTTGAAATCCAGAGCATCAATGCTCTTCTGGTGTTCACCATAGCCACAGCAGA 3088  
 DB 3 GGAACGTTGAAATCCAGAGCATCAATGCTCTTCTGGTGTTCACCATAGCCACAGCAGA 62

QY 3089 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGCAGGATTTAAGAAATGTAACCTAT 3148  
 DB 63 TGTCTTAATCTTCCGAGATCTAGTTTTCAGCAAGCAGGATTTAAGAAATGTAACCTAT 122

QY 3149 CTTATGCTGTATGAAGCAACATGAATCAATGCTGTATAGTGCTTTTAACTGTAA 3208  
 DB 123 CTTATGCTGTATGAAGCAACATGAATCAATGCTGTATAGTGCTTTTAACTGTAA 182

QY 3209 TTTTGTGAAGCTATCTTTTATGCAATAAATAATTTGAACATTTTACATGTTATATTT 3268  
 DB 183 TTTTGTGAAGCTATCTTTTATGCAATAAATAATTTGAACATTTTACATGTTATATTT 242

QY 3269 TTAATCAGTTTACTCAAGTGTGATTTATACAGAAAATGTAACCATGTAAGGCTAGA 3328  
 DB 243 TTAATCAGTTTACTCAAGTGTGATTTATACAGAAAATGTAACCATGTAAGGCTAGA 302

QY 3329 GTTATAAGAAATTTTGTCAAATGTATTCACCCATGTAGTCACCTCCTTATGAAGAGCAGA 3388  
 DB 303 GTTATAAGAAATTTTGTCAAATGTATTCACCCATGTAGTCACCTCCTTATGAAGAGCAGA 362

QY 3389 ACAGTACATCTCCAGAAAGTTCCACAGTGCTCTTTTCCCTGAGTTTCCACAGTCCT 3448  
 DB 363 ACAGTACATCTCCAGAAAGTTCCACAGTGCTCTTTTCCCTGAGTTTCCACAGTCCT 422

QY 3449 GGCAACCAATGATCTGCTGTATATTAATTAACCTGTAGATATTTGTAGCAATGACC 3508  
 DB 423 GGCAACCAATGATCTGCTGTATATTAATTAACCTGTAGATATTTGTAGCAATGACC 482

QY 3509 CTTTCCATATTTATTTGTGTGTGAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGAT 3568  
 DB 483 CTTTCCATATTTATTTGTGTGTGAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGAT 542

QY 3569 TCATCTATGTTTAAATGTTTATCAGTAGTGTGACATCTTACTTGTCTCAGCATATCACA 3628  
 DB 543 TCATCTATGTTTAAATGTTTATCAGTAGTGTGACATCTTACTTGTCTCAGCATATCACA 3628

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 18, 2004, 13:34:28 ; Search time 10878 Seconds  
(without alignments)  
12602.158 Million cell updates/sec

Title: US-10-005-907-1  
Perfect score: 3762  
Sequence: 1 gagaaacagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gss1:  
9: gb\_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	696.8	18.5	749	5	BU852381
2	695.8	18.5	782	5	BQ441475
3	694.2	18.5	704	8	AQ390121
4	692.4	18.4	785	5	BU171872
5	673	17.9	741	5	BU155309
6	628.6	16.7	786	5	BU155429
7	612.2	16.3	728	6	CD639433
8	570.4	15.2	643	9	AG182348
9	533.8	14.2	549	2	BF434786
10	506	13.5	527	2	BF242113
11	502	13.3	809	2	BF240893
12	500.8	13.3	504	2	AW655849
13	499.6	13.3	757	8	BZ600841
14	487.6	13.0	848	2	BF243193
15	446.6	11.9	456	1	AI766257
16	410.8	10.9	549	4	BG461807
17	383.2	10.2	416	1	AA447527
18	328	8.7	386	8	AQ006627
19	260	6.9	920	8	AQ743145
20	232.4	6.2	439	8	B59695
21	226	6.0	1045	4	BM925124
22	221.8	5.9	437	8	AQ109243
23	198	5.3	475	8	B81885
24	196.2	5.2	611	6	CD704731

BM983330 UI-CP-DU1  
AQ591876 HS\_5436.A  
CR546941 DKFZp470G  
BI857056 603383626  
BU182600 AGENCOURT  
EG680170 602626836  
AG084541 Pan trogl  
BE779049 601464810  
AA618471 mn27h08.8  
AQ827545 HS\_5304.A  
BM467056 AGENCOURT  
T74524 YC83C08.r1  
BE935507 601311094  
B86762 RPC111-24F1  
AQ740505 HS\_5504.A  
BU596567 AGENCOURT  
BM989036 UI-H-DF0-  
BX480744 DKFZp686P  
BG942531 ax26d03.x  
CK902534 im66g10.x  
BU078278 im66g10.y

ALIGNMENTS

RESULT 1  
BU852381  
LOCUS BU852381 749 bp mRNA linear EST 16-OCT-2002  
DEFINITION AGENCOURT\_10402421 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6617922  
5', mRNA Sequence.  
ACCESSION BU852381  
VERSION BU852381.1 GI:24037344  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 749)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHC2863 row: p column: 18  
High quality sequence stop: 659.  
Location/Qualifiers  
1. 749  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6617922"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NIH MGC 82"  
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1:  
sf11 (ggcgccctggcc); Site 2: sf11 (ggccattatggcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size  
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained  
inserts by PCR. This library was enriched for full-length  
clones and was constructed by Clontech Laboratories (Palo  
Alto, CA)."

ORIGIN

Page Blank (uspto)

Fri Nov 19 05:43:42 2004

Db 29911 AATCCAGCTACTCGGAGGCTGAGTCTAGGAGATCACTTGACCCAGAGGCGCAGAGTT 29970  
Qy 2534 GAACTGAGCCAGGATCTTGGCCACACATTCAGCTTGGGCAACAGAGTGAGACCCCTGTCT 2593  
Db 29971 GCAATGAGCCAGGATCACTGCTACTTCCAGCCTGGTGACAGAAATGAGACTCAATCT 30030  
Qy 2594 CAAAAAAGGTTTCAGATTTTGGAGCAATTT 2629  
Db 30031 CAAAAAAGGTTTCAGATTTTGGAGCAATTT 30066

RESULT 15  
US-08-429-742-1  
; Sequence 1, Application US/08429742  
; Patent No. 5686257  
; GENERAL INFORMATION:  
; APPLICANT: Kennedy, Jacqueline  
; APPLICANT: Bazan, J. Fernando  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: PURIFIED MAMMALIAN T CELL ANTIGENS AND  
; TITLE OF INVENTION: RELATED REAGENTS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DNAX Research Institute  
; STREET: 901 California Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/429,742  
; FILING DATE: 26-APR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ching, Egwin P.  
; REGISTRATION NUMBER: 34,090  
; REFERENCE/DOCKET NUMBER: DX0505  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-852-9196  
; TELEFAX: 415-496-1200  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2477 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 38..1219  
; US-08-429-742-1

Query Match 4.9%; Score 183.8; DB 1; Length 2477;  
Best Local Similarity 77.0%; Pred. No. 5.4e-28;  
Matches 224; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
Qy 2316 CTGGTATGAGTTTCAGGAGGCTGAGTCTTGGGAGGCTGAGTCTTGGGAGGCTGAGTCT 2375  
Db 1676 CAGGCGGGTGGCTCATGCTGTAATCTAGCACTTTGGGAGGCTGAGTCTTGGGAGGCT 1735  
Qy 2376 CTTGAGCCAGGAGTTTGGGAGGCTGAGTCTTGGGAGGCTGAGTCTTGGGAGGCT 2435  
Db 1736 CTTGAGGTCAGGAGTTTGGGAGGCTGAGTCTTGGGAGGCTGAGTCTTGGGAGGCT 1795  
Qy 2436 AATTAATAAATAGCCAGGCTGAGTCTTGGGAGGCTGAGTCTTGGGAGGCTGAGTCT 2495  
Db 1796 AATGCAAAATAGCCAGGCTGAGTCTTGGGAGGCTGAGTCTTGGGAGGCT 1855

Qy 2496 GAGGTAGTAGGATTTTGTGAGACTGGGAGGTTGAGGCTGAACTGAGCCAGGATTTGCCA 2555  
Db 1856 GAGGTGGAGATCGCTTTGAACCTGGGAGGTTGGAGATTGCGAGTGAATCTCAACA 1915  
Qy 2556 CCATATCCAGCTTTGGGCAACAGAGTGAGACCTGTCTTCAAAAAAAGAAA 2606  
Db 1916 CTGAACTCCAGCTGGGCGACAGAGGGAGACTCTGTCTCTCAAAAAAAGAAA 1966

Search completed: November 18, 2004, 21:31:25  
Job time : 291 secs

; Sequence 21, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Krommal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Teuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,394A  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 21:  
; LENGTH: 246240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..246240  
; OTHER INFORMATION: /note= "HLA-H. CONTIG"  
; US-08-724-394A-21

Query Match 4.9%; Score 184; DB 2; Length 246240;  
Best Local Similarity 71.7%; Pred. No. 1.6e-27;  
Matches 241; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
QY 2294 TCAAAAAGTTTCAGATGTTAAGCTGGTGTGATGCGAGTTTCATGCCAGTGCAGTACTTTG 2353  
Db 29731 TAAAAAAGTTTCAGATGTTAAGCTGGTGTGATGCGAGTTTCATGCCAGTGCAGTACTTTG 29790  
QY 2354 GGAAGCCACAGAGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGCACACACA 2413  
Db 29791 GGAGGCTAAGCGAGGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGCACACACA 2413  
QY 2414 GTGAGACCTCTGTTTCTACAAATTAATAAATAATAGCCAGGTGTGGTGGTGCACACTGT 2473  
Db 29851 GTGAACCCCATCTCTACTAAAATAACAAAATTAACCCAGCATGGTGTGATGCTGT 29910  
QY 2474 AGTCCAGGTACTCAGAGGCTGAGGTAGTGTAGGATTTTGTAGACTGGAGGTGTAGGCT 2533  
Db 29911 AATCCAGTACTCGGAGGCTGAGTCAGAGAGTCACTTGAACCCAGGAGGAGGT 29970  
QY 2534 GAATGAGCCAGATCTTGCCACACATTCAGCTTGGGCAACAGAGTGAGACCTGTCT 2593  
Db 29971 GCAATGAGCCAGATCATGTCTACTTCCAGCTGGGTGACAGAGTCACTCAATCT 30030

QY 2594 CAAAAAAGTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTT 2629  
Db 30031 CAAAAAAGTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTT 30066  
RESULT 14  
US-08-724-394A-22  
; Sequence 22, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Krommal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Teuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,394A  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 22:  
; LENGTH: 246240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..246240  
; OTHER INFORMATION: /note= "HLA-H. CONTIG"  
; US-08-724-394A-22

Query Match 4.9%; Score 184; DB 2; Length 246240;  
Best Local Similarity 71.7%; Pred. No. 1.6e-27;  
Matches 241; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
QY 2294 TCAAAAAGTTTCAGATGTTAAGCTGGTGTGATGCGAGTTTCATGCCAGTGCAGTACTTTG 2353  
Db 29731 TAAAAAAGTTTCAGATGTTAAGCTGGTGTGATGCGAGTTTCATGCCAGTGCAGTACTTTG 29790  
QY 2354 GGAAGCCACAGAGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGCACACACA 2413  
Db 29791 GGAGGCTAAGCGAGGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGCACACACA 2413  
QY 2414 GTGAGACCTCTGTTTCTACAAATTAATAAATAATAGCCAGGTGTGGTGGTGCACACTGT 2473  
Db 29851 GTGAACCCCATCTCTACTAAAATAACAAAATTAACCCAGGATGGTGTGATGCTGT 29910  
QY 2474 AGTCCAGGTACTCAGAGGCTGAGGTAGTGTAGGATTTTGTAGACTGGAGGTGTAGGCT 2533





```

Db 143 CTCAGGAGTTCAAGAACACCTCGCCAAACATGGTGAACCCCATCTCTACTAAATAACA 202
QY 2442 AAAATTAGCCAGGTGCTGGTGGTGACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTA 2501
Db 203 AAAATTAGTGGGCATGGTGGTGGGAGGCTGTAAATCCCAGTTACTTTGGGAGGCTAAGGTA 262
QY 2502 GTAGGATTTGTTTGAAGCTGGGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCACCAAT 2561
Db 263 GCAGAAATTACTTGAATCTGGGAGGAGGTTGAGTGAGCCAGATTTGCCCACTGCAC 322
QY 2562 TCCAGCTGGGCAACAGAGTGAGACCCCTGTCTCAAAAAAAGGTTTTCAGATTT 2619
Db 323 TCCAGCCTGGGTGACAGAGTCAGACTCTGTCTCAAAAAAAGGTTTTCAGATTT 380

```

## RESULT 9

```

US-09-754-250-3/c
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

```

```

Query Match 4.9%; Score 184.2; DB 3; Length 111282;
Best Local Similarity 76.8%; Pred. No. 1.2e-27;
Matches 225; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 2316 CTGCTGATGCAGTTTCATGCCAGTGCATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCT 2375
Db 14773 CGGTGCGAGTGGCTATGCTCTGTAATCCAGTACTTTGGGAGGCCAAGGTGGGTGATCA 14714
QY 2376 CTTGAGCCCGAGGTTTGGGCCAGACTGCACACACAGTGAGACCTCTGTTTCTACAAAT 2435
Db 14713 TTTGAGGTCAGGAGTTCAAGACACAGCCCTGGCCAACTGGTGAAACCCCATCTCTACTAA 14654
QY 2436 AATTAAAAAATTAGCCAGGTGGTGGTGACACCTGTAGTCCAGGTACTCAGGAGGCT 2495
Db 14653 AATACAAAAAATTATGCAAGGTGGTGGTGATCGCTGTAGTCCAGGTACTCAGGAGGCT 14594
QY 2496 GAGGTAGTAGGATTTGAGACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCA 2555
Db 14593 GAGGAGGAGAAATAGCTTGAGCTTGGATCGGAGGTTGGGTGAGCCGAGATCAGCCA 14534
QY 2556 CCACATTCAGCTTGGGCAACAGAGTGAGACCCCTGTCTCAAAAAAAGGTTTTCAGATTT 2608
Db 14533 CTGCACTCCAGTCTGGGTGAGAGTGAGACCCCTGTCTCAAAACTAACTAA 14481

```

## RESULT 10

```

US-08-370-975B-6/c
; Sequence 6, Application US/08370975B
; Patent No. 5622851
; GENERAL INFORMATION:
; APPLICANT: Maley, Frank
; APPLICANT: Maley, Gladys F.
; APPLICANT: Weiner, Karen X.B.
; TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene

```

```

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,975B
; FILING DATE: 10-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20894/80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4q35
US-08-370-975B-6

```

```

Query Match 4.9%; Score 184; DB 1; Length 20303;
Best Local Similarity 74.7%; Pred. No. 8.6e-28;
Matches 257; Conservative 0; Mismatches 85; Indels 2; Gaps 2;

```

```

QY 2280 TTTCAAGTCAGTCTCAAAAGTTTCAGATGTTAAG-CTGGTGCATGCAGTTTCATGCCAGT 2338
Db 14915 TGTCATGCTGCTGTGCAAAAGTTTCAGATTTGGGCCAGGTGCAGTAGTTCATGCCCTGT 14856
QY 2339 GATCCAGGACTTTTGGGAAGCCAGACAGGTGGATCTCTTGAGCCAGGAGTTTGAGGCC 2398
Db 14855 AATCCAGGACTTTTGGGAGGCCAAGTGGGGAGATCGTTTGAGCTCAAGAGTTCAAGAAC 14796
QY 2399 AGACTGCCAACACACAGTGAGACCTCGTTTCTACAAATATTAAAAA-TTAGCCAGGTGT 2457
Db 14795 AGCTTAGGCAACATAGTGACAGCTCATCTCTACTGAAAAATAAAAAAAGTTAGCCAGTAT 14736
QY 2458 GGTGGTGACACCTGTAGTCCCAGGTACTCAGGAGGCTCAGGTAGTAGGATTTGTTGAGA 2517
Db 14735 GGTGGCGTGCACCTGTAGTCTCTAGTACTTTGGAGGCTCAGGTGGAGGAGCTGCTTGAGC 14676
QY 2518 CTGGGAGGTTGAGGCTGAACTGAGCAGGATCTTGCCACCACTTCAGCTTGGGCAACA 2577
Db 14675 CCAGGAGGTCAGAGGCTGAGTGTGTTTGGCACCCTGTACTCCAGTCTCGGCAACA 14616
QY 2578 GAGTGAGACCCCTGTCTCAAAAAAAGGTTTTCAGATTTTG 2621
Db 14615 GAGCAAGACCCCATCTCAGAAAAAAGGTTTTCAGATTTTG 14572

```

## RESULT 11

```

US-08-370-975B-1/c
; Sequence 1, Application US/08370975B
; Patent No. 5622851
; GENERAL INFORMATION:
; APPLICANT: Maley, Frank
; APPLICANT: Maley, Gladys F.
; APPLICANT: Weiner, Karen X.B.
; TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene

```

LOCATION: 15242..15260  
OTHER INFORMATION: 17-41-250.mis complement  
NAME/KEY: primer bind  
LOCATION: 42199..42217  
OTHER INFORMATION: 20-841-149.mis  
NAME/KEY: primer bind  
LOCATION: 42219..42237  
OTHER INFORMATION: 20-841-149.mis complement  
NAME/KEY: primer bind  
LOCATION: 45423..45441  
OTHER INFORMATION: 20-842-115.mis  
NAME/KEY: primer bind  
LOCATION: 45443..45461  
OTHER INFORMATION: 20-842-115.mis complement  
NAME/KEY: primer bind  
LOCATION: 77039..77057  
OTHER INFORMATION: 20-853-415.mis  
NAME/KEY: primer bind  
LOCATION: 77059..77077  
OTHER INFORMATION: 20-853-415.mis complement  
NAME/KEY: misc binding  
LOCATION: 1227..1251  
OTHER INFORMATION: 20-828-311.probe  
NAME/KEY: misc binding  
LOCATION: 12335..12359  
OTHER INFORMATION: 17-42-319.probe  
NAME/KEY: misc binding  
LOCATION: 15229..15253  
OTHER INFORMATION: 17-41-250.probe  
NAME/KEY: misc binding  
LOCATION: 42205..42230  
OTHER INFORMATION: 20-841-149.probe  
NAME/KEY: misc binding  
LOCATION: 45430..45454  
OTHER INFORMATION: 20-842-115.probe  
NAME/KEY: misc binding  
LOCATION: 77045..77070  
OTHER INFORMATION: 20-853-415.probe  
US-09-750-580-1

Query Match 4.9%; Score 185; DB 4; Length 81001;  
Best Local Similarity 77.5%; Pred. No. 7.7e-28;  
Matches 224; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 2318 GGTGATGCGTTCATGCCAGTGTCCGAGTACTTGGAGGCGCAACAGAGGTGATCTCT 2377  
DB 69837 GGGCCAGTGGCTCATGCTATGATCCTAGTGTCTTGGGAGGCTGAGGTGGGTGCT 69896  
QY 2378 TGAGCCAGGAGTTTGGAGCCAGTGCACACAGAGTGAAGCTCTTCTACAATAA 2437  
DB 69897 TGAGCCAGGAGTTTGGAGCCAGTGCACACAGAGTGAAGCTCTTCTACAATAA 69956  
QY 2438 TTAATAAATAGCCAGGTTGGTGGTGCACACCTGTAGTCCAGGTAAGTCCAGGCTGA 2497  
DB 69957 TTAATAAATAGTGGCGATGGTGGCCACCTGTAGTCCAGGTAAGTCCAGGCTGA 70016  
QY 2498 GGTAGTAGATTTTGGAGCTGGAGGTTGAGCTGAATGAGCCAGGATCTTGCACC 2557  
DB 70017 GGTGAAGGATCATCTTGAATCTGGAGGTTGAGCTGCACTGAGCTGTGATTGAGCCACT 70076  
QY 2558 ACATTCAGGTTGGGCAACAGAGTGAAGCCCTGTCTCAAAAAAATAA 2606  
DB 70077 GCATCCAGCCGGTGACAGAGTGAAGCTGTCTCTCAAAAAAATAA 70125

RESULT 7  
US-09-497-855A-40  
Sequence 40, Application US/09497855A  
Patent No. 6605432  
GENERAL INFORMATION:  
APPLICANT: Huang, Tim  
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION  
FILE REFERENCE: UWO1523

CURRENT APPLICATION NUMBER: US/09/497,855A  
CURRENT FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/120,592  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: 60/118,760  
PRIOR FILING DATE: 1999-02-05  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 40  
LENGTH: 161652  
TYPE: DNA  
ORGANISM: Homo sapiens;  
US-09-497-855A-40  
Query Match 4.9%; Score 184.8; DB 4; Length 161652;  
Best Local Similarity 77.1%; Pred. No. 1e-27;  
Matches 225; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
QY 2327 GTTCATGCCAGTGTCCGAGTACTTTGGGAAGCCAGAGAGGTGATCTCTTGGAGCCAG 2386  
DB 154102 GCTCAGCGCTATAATCTCAGCAGCTTTGGAGGCGGAGGAGGTGGATCATTGAGGTCAG 154161  
QY 2387 GAGTTTGGAGCCAGACTGCACAACAGTGCAGCTCTCTCAAAATAATTAATAAAT 2446  
DB 154162 GAAATTTGAGACCAAGTCTGGCCAATACGGTGAACCTGTCTTACTAAATAATACAAAAT 154221  
QY 2447 TAGCCAGGTGTGGTGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGG 2506  
DB 154222 TACCAGGTGTGGTGTCCACACCTGTAGTCTTAAGTACTCAGGAGGCTGAGGCGAGAGA 154281  
QY 2507 ATTGTTGAGACTGGGAGGTTGAGGCTGAAGTGCAGGAGTGTGAGGAGGCTGAGGTAGTAGG 2566  
DB 154282 ATTGTTGAACTGGGAAGCAGAGGTTGAGTGCAGGAGGAGTGCAGGAGTGCAGTCCAG 154341  
QY 2567 CTGGGCAACAGAGTGCAGCCCTGTCTCAAAAAAATAAAGTTTCAGATT 2618  
DB 154342 CCTGAGTGCAAGCAAGACTCTGTCTCAAAAAAATAAAGTTTCAGATT 154393

RESULT 8  
US-09-621-976-13959  
Sequence 13959, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 13959  
LENGTH: 399  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-13959

Query Match 4.9%; Score 184.4; DB 4; Length 399;  
Best Local Similarity 69.3%; Pred. No. 2.6e-28;  
Matches 248; Conservative 2; Mismatches 108; Indels 0; Gaps 0;  
QY 2262 AAGCAGAGTGTCACTATTTCAGTCAGTGCCTCAAAAGCTTTCAGATGTTAAGCTGGTG 2321  
DB 23 AAGGAAAAAAGAAATTTTGGATTGTTAGTCCATAAATGTTCTGGTGGCCGCG 82  
QY 2322 ATGAGTTTCATGCCAGTGTCCGAGTACTTTGGGAAGCCAGAGAGGTGATCTTTGAG 2381  
DB 83 TGGTGGCTCATGCTGTAACTCTAGCACTTTGGAGGCTGAGGAGGTTGATCATTTGAG 142  
QY 2382 CCAGGAGTTTGGAGGCGAGACTGCACAACAGTGCAGCTCGTTCTTACAAATAATAA 2441

```
QY 2357 AGCCAGACAGGTGATCTCTTGGCCAGGAGTTTGGCCAGACTGCACACACAGTG 2416
Db 594 AGCCAGGTGGTGGATCATCTAGGTGAGGATTTAAGACAGGCTGGCCACATGGCA 653
QY 2417 AGACCTGTTTCTACAAATAATTAATAAATTAGCCAGGTGTGGTGGTGCACACCTGTAGT 2476
Db 654 AAACCTCGTCTCTACTAAATAATACAAAAATAGCCAGGCTTGGTGGTGCACACCTGTAT 713
QY 2477 CCCAGTACTCAGGAGGCTGAGGTAGTAGTATTGTTAGACTGGAGGTTGAGGCTGAA 2536
Db 714 CCCAGTACTCAGGAGGCTGAGGCAAGAGATCGCTTGAACCTGGAGGTGGAGGTTGCA 773
QY 2537 CTGAGCCAGGATCTTCCACACACATTCAGCTTGGGCAACAGAGTGAGACCTGTCTCAA 2596
Db 774 GTGAGCCGAGATCAGCCACTGCATCTAGCTTGGGCAACAGAGCAAGACTCTATCTCAA 833
QY 2597 AAAAAAAGTTTCAGATTTG 2621
Db 834 AAAAAAAGAGAGTTTCTTCTG 858

RESULT 6
US-09-750-580-1
; Sequence 1, Application US/09750580
; Patent No. 6455280
; GENERAL INFORMATION:
; APPLICANT: Yen, Frances
; APPLICANT: Denison, Blake
; APPLICANT: Bour, Barbara
; APPLICANT: Bihain, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Ebbets-Read, Dana
; APPLICANT: Salter-Cid, Luisa
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
; FILE REFERENCE: 89.US2.CIP
; CURRENT APPLICATION NUMBER: US/09/750,580
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 09/599,362
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB00/0101
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB99/02058
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 49/469/099
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 60/141,032
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 81001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10946..12946
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 12947..12958
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 13470..13526
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 13641..13752
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 14271..15968
; OTHER INFORMATION: exon 4
; NAME/KEY: misc feature
; LOCATION: 15969..17969
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1239
; OTHER INFORMATION: 20-828-311 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12347
; OTHER INFORMATION: 17-42-319 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 15241
; OTHER INFORMATION: 17-41-250 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 42218
; OTHER INFORMATION: 20-841-149 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 45442
; OTHER INFORMATION: 20-842-115 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 77058
; OTHER INFORMATION: 20-853-415 : polymorphic base C or T
; NAME/KEY: primer_bind
; LOCATION: 929..929
; OTHER INFORMATION: 20-828.pu
; NAME/KEY: primer_bind
; LOCATION: 1357..1377
; OTHER INFORMATION: 20-828.rp complement
; NAME/KEY: primer_bind
; LOCATION: 12029..12050
; OTHER INFORMATION: 17-42.pu
; NAME/KEY: primer_bind
; LOCATION: 12581..12603
; OTHER INFORMATION: 17-42.rp complement
; NAME/KEY: primer_bind
; LOCATION: 14992..15012
; OTHER INFORMATION: 17-41.pu
; NAME/KEY: primer_bind
; LOCATION: 15460..15482
; OTHER INFORMATION: 17-41.rp complement
; NAME/KEY: primer_bind
; LOCATION: 42070..42090
; OTHER INFORMATION: 20-841.pu
; NAME/KEY: primer_bind
; LOCATION: 42572..42591
; OTHER INFORMATION: 20-841.rp complement
; NAME/KEY: primer_bind
; LOCATION: 45328..45347
; OTHER INFORMATION: 20-842.pu
; NAME/KEY: primer_bind
; LOCATION: 45863..45883
; OTHER INFORMATION: 20-842.rp complement
; NAME/KEY: primer_bind
; LOCATION: 76644..76664
; OTHER INFORMATION: 20-853.pu
; NAME/KEY: primer_bind
; LOCATION: 77166..77185
; OTHER INFORMATION: 20-853.rp complement
; NAME/KEY: primer_bind
; LOCATION: 1220..1238
; OTHER INFORMATION: 20-828-311.mis
; NAME/KEY: primer_bind
; LOCATION: 1240..1258
; OTHER INFORMATION: 20-828-311.mis complement
; NAME/KEY: primer_bind
; LOCATION: 12328..12346
; OTHER INFORMATION: 17-42-319.mis
; NAME/KEY: primer_bind
; LOCATION: 12348..12366
; OTHER INFORMATION: 17-42-319.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15222..15240
; OTHER INFORMATION: 17-41-250.mis
; NAME/KEY: primer_bind
```

```

; OTHER INFORMATION: exon 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T

```

```

; FEATURE:
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6019
; OTHER INFORMATION: 10-346-141 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6141
; OTHER INFORMATION: 10-346-263 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6183
; OTHER INFORMATION: 10-346-305 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6338
; OTHER INFORMATION: 10-347-74 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6375
; OTHER INFORMATION: 10-347-111 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6429
; OTHER INFORMATION: 10-347-165 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6467
; OTHER INFORMATION: 10-347-203 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6484
; OTHER INFORMATION: 10-347-220 : polymorphic base A or G

```

```

Query Match      4.9%; Score 185.8; DB 4; Length 20674;
Best local Similarity 73.2%; Pred. No. 3.7e-28;
Matches 238; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy      2297 AAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTCATGCCAGTATCCGAGTACTTTGGGA 2356
Db      534 AAAAATCTGAAGAGCTGGTGGCATGTGGCTCACGCTGTATATCCAGCACTTTGGGA 593

```

OTHER INFORMATION: 10-347-165 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 6467  
OTHER INFORMATION: 10-347-203 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6484  
OTHER INFORMATION: 10-347-220 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6534  
OTHER INFORMATION: 10-347-271 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 6611  
OTHER INFORMATION: 10-347-348 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 7668  
OTHER INFORMATION: 10-348-391 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 8608  
OTHER INFORMATION: 10-349-47 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 8658  
OTHER INFORMATION: 10-349-97 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 8703  
OTHER INFORMATION: 10-349-142 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 8777  
OTHER INFORMATION: 10-349-216 : deletion of C/T  
NAME/KEY: allele  
LOCATION: 8785  
OTHER INFORMATION: 10-349-224 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 8926  
OTHER INFORMATION: 10-349-368 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 12171  
OTHER INFORMATION: 10-350-72 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 12429  
OTHER INFORMATION: 10-350-332 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 13341  
OTHER INFORMATION: 10-507-170 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 13492  
OTHER INFORMATION: 10-507-321 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 13524  
OTHER INFORMATION: 10-507-353 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 13535

Query Match 4.9%; Score 185.8; DB 4; Length 20674;  
Best Local Similarity 73.2%; Pred. No. 3.7e-28;  
Matches 238; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
QY 2297 AAAAGTTTCAGATGTTAAGCTGTGTGATGCGATTCATGCCAGTCATCCGAGTACTTTGGGA 2356  
DB 534 AAAAATCTGAAAGCTGGCTGGCGATGTGGCTCAGCCCTGTAAATCCAGCACTTTGGGA 593  
QY 2357 AGCCAGACAGGTGGATCTCTTGAGCCCGAGGAGTTTGAGCCAGACTGACACACACAGTG 2416  
DB 594 AGCCGAGGTGGTGGATCATCTGAGGTCAGAGTTTAAAGACCCAGCTGGCCACATGGCA 653  
QY 2417 AGACCTGTTTCTACAAATTAATTAATAATAGCCAGGTGTGGTGACACCTGTAGT 2476  
DB 654 AAACCTGCTCTACTATAAATAACAAAACATAGCCAGGTTGGTGGTGACACCTGTAT 713  
QY 2477 CCCAGGTACTCAGGAGCTGAGGTAGTAGGATTTTGACCTGGGAGGTGGAGCTGAA 2536  
DB 714 CCAGCTACTCAGAGAGCTGAGCAGAGAAATCGCTTGAACCTGGGAGGTGGAGTTGCA 773  
QY 2537 CTGAGCCAGGATCTTGCCACCAATTCAGCTTTGGGCAACAGAGTGAGACCCCTGTCTCAA 2596

Db 774 GTGAGCCGAGATCAGCCACTGCACTTAGCTGGCAACAGAGCAAGACTCTATCTCAA 833  
QY 2597 AAAAAAAGTTTCAGATTTTG 2621  
DB 834 AAAAAAAGAGAGTTTCTCTG 858  
RESULT 5  
US-10-170-097-651  
; Sequence 651, Application US/10170097  
; Patent No. 6794143  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM  
; FILE REFERENCE: GEN-T114XC2D1  
; CURRENT APPLICATION NUMBER: US/10/170,097  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US 09/641,638  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 09/502,330  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: US 60/133,200  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: US 09/275,267  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: US 60/119,917  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 1304  
; SOFTWARE: Patent.pm  
; SEQ ID NO 651  
; LENGTH: 20674  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; LOCATION: 1123..3123  
; OTHER INFORMATION: 5'regulatory region  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 3124..3297  
; OTHER INFORMATION: exon 1  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 3871..4072  
; OTHER INFORMATION: exon 2  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 5552..5633  
; OTHER INFORMATION: exon 3  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 5758..5880  
; OTHER INFORMATION: exon 4  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 5996..6099  
; OTHER INFORMATION: exon 5  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 6349..6509  
; OTHER INFORMATION: exon 6  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 7379..7522  
; OTHER INFORMATION: exon 7  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 8645..8854

```
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051c1p1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; NAME/KEY: misc.feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTTT
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6019
; OTHER INFORMATION: 10-346-141 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6141
; OTHER INFORMATION: 10-346-263 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 6183
; OTHER INFORMATION: 10-346-305 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 6338
; OTHER INFORMATION: 10-347-74 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6375
; OTHER INFORMATION: 10-347-111 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 6429
```



## RESULT 2

US-10-140-002-383  
; Sequence 383, Application US/10140002  
; Patent No. 6725730

## ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C59  
; CURRENT APPLICATION NUMBER: US/10/140,002  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 383  
; LENGTH: 2336  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 1620, 1673  
; OTHER INFORMATION: unknown base

US-10-140-002-383

Query Match 5.0%; Score 186.6; DB 4; Length 2336;  
Best Local Similarity 73.6%; Pred. No. 1.5e-28;  
Matches 251; Conservative 0; Mismatches 89; Indels 1; Gaps 1;

Qy	2276	ACTATTTCAAGTCAGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGATGAGTTTCATGCC	2335
Db	1955	ACATTTTCCCATTTCTGTCTCATCAAAAAGTTTCAGATGTTAAGCTGGTGATGAGTTTCATGCC	2014
Qy	2336	AGTGATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCTTTGAGCCCGAGGTTTGAG	2395
Db	2015	TGTAATCCAGCACTTTGGG-GGCCAAGAGGGTGGATCACTTGAGATCAGGAGTTCAAG	2073
Qy	2396	GCAGACTGCACAAACACAGTGAGACCTCGTTTCTACAAAATAATTAATAAATTAGCCAGGT	2455
Db	2074	ACCAGCTGGCCCAACATGGTGAACCTTGTTCTACTAAAAATACAAAATTAGCCAGGC	2133
Qy	2456	GTGTTGTGCACACCTGTAGTCCAGGTACTCAGGAGCTGAGGTAGTAGATTGTTTGA	2515
Db	2134	GTGTTGTGCACACCTGTAGTCCAGGTACTCAGGAGCTGAGCAGGAGATTGCTTGA	2193
Qy	2516	GACTGGGAGGTGAGGCTGAACCTGAGCCAGGATCTTCCACCAATTCAGACTTGGCAA	2575
Db	2194	ACCGGAGGCGGAGGTTCAGTGAGCCAGATTTGTCCTCAGCTCCAGCTGGGTGA	2253
Qy	2576	CAGAGTGAACCTGTCTCAAAAAAAGTTTTCAGA	2616
Db	2254	CAGAGCAAGACTCCATCTCAAAAAAAGTTTTCAGA	2294

## RESULT 3

US-09-620-312D-967  
; Sequence 967, Application US/09620312D

; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John fillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: PC\_FL\_genes Version 1.0  
; SEQ ID NO 967  
; LENGTH: 1426  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (41)..(961)  
; OTHER INFORMATION: unknown base

US-09-620-312D-967

Query Match 4.9%; Score 186.2; DB 4; Length 1426;  
Best Local Similarity 75.9%; Pred. No. 1.5e-28;  
Matches 230; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy	2306	AGATGTTAAGCTGGTGATGATCCAGTTCATGCCAGTATCCGAGTACTTTGGGAAGCCAGAC	2365
Db	974	AGAAGATGGCGGCGGTGGCTCATGCTGTAATTCAGCACTTTGGGAGGCTAAGGC	1033
Qy	2366	AGGTGGATCTTTGAGCCCGAGGTTGAGCCAGACTGCACACACAGTGAAGCTCGT	2425
Db	1034	GGGAGGATGATTTGAGCCCGAGGTTGAGCCAGCTGGCCAACTGGTGAACCTCTGT	1093
Qy	2426	TTCTCAAAATAATTAATAAATTAGCCAGGTTGGTGGTGACACACCTGTAGTCCAGGTAC	2485
Db	1094	CTCTACTAAAATACAAAATTAAGCTGGTGGTGACACACCTGTATCCAGCTAC	1153
Qy	2486	TCAGGAGGCTGAGGTAGTAGATTGTTGAGACTGGGAGGTTGAGGCTGAACCTGAGCCAG	2545
Db	1154	TCGTGAGGCTGAGGAGGAGGAATCACTTGAACCTGGAGGAGGTTGCAATGAGCCGA	1213
Qy	2546	GATCTTGGCCACCACTTCCAGCTTGGCCAAACAGAGTGAGACCTGTCTCAAAAAA	2605
Db	1214	GATCACCACTACACTCCAGGCTAGGCAACAGAGCAAGACTCTGTCTCAAAAAA	1273
Qy	2606	AAA	2608
Db	1274	AAA	1276

## RESULT 4

US-09-641-638-651  
; Sequence 651, Application US/09641638  
; Patent No. 6432648  
; GENERAL INFORMATION:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 13:35:33 ; Search time 288 Seconds  
(without alignments)  
9284.680 Million cell updates/sec

Title: US-10-005-907-1  
Perfect score: 3762  
Sequence: 1 gagaaacagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187.8	5.0	1332	4	US-09-584-568C-1
2	186.6	5.0	2336	4	US-10-140-002-383
3	186.2	4.9	1426	4	US-09-620-312D-967
4	185.8	4.9	20674	4	US-09-641-638-651
5	185.8	4.9	20674	4	US-10-170-097-651
6	185.4	4.9	81001	4	US-09-750-580-1
7	184.8	4.9	161652	4	US-09-497-855A-40
8	184.4	4.9	399	4	US-09-621-976-13959
9	184.2	4.9	111282	3	US-09-754-250-3
10	184.2	4.9	20303	1	US-08-370-975B-6
11	184.2	4.9	26764	1	US-08-370-975B-1
12	184.2	4.9	246240	2	US-08-724-394A-20
13	184.2	4.9	246240	2	US-08-724-394A-21
14	184.2	4.9	246240	2	US-08-724-394A-22
15	183.8	4.9	2477	1	US-08-429-742-1
16	183.4	4.9	2932	4	US-09-016-434-1419
17	183.4	4.9	2932	4	US-09-054-272-5
18	183.4	4.9	6669	3	US-09-212-971-5
19	183.4	4.9	6669	3	US-08-800-929A-5
20	183.4	4.9	6669	3	US-09-617-053A-5
21	183.4	4.9	6669	4	US-09-672-717-230
22	182.8	4.9	1430	4	US-09-489-847-39
23	182.6	4.9	48763	4	US-09-916-204-3
24	182.6	4.9	48763	4	US-10-282-048-3
25	182.6	4.9	174493	4	US-09-804-471A-3
26	182.6	4.9	174493	4	US-10-238-709-3
27	182.2	4.8	1855	4	US-09-023-655-1380

c	28	182	4.8	786431	4	US-09-751-389-3	Sequence 3, Appli
	29	181.8	4.8	4768	4	US-09-526-193A-16	Sequence 16, Appli
	30	181.4	4.8	107820	4	US-09-792-616-1	Sequence 1, Appli
	31	180.6	4.8	1504	4	US-10-067-422-8	Sequence 8, Appli
	32	180.6	4.8	1519	4	US-09-716-129-45	Sequence 45, Appli
	33	180.6	4.8	44848	4	US-09-435-739-42	Sequence 42, Appli
	34	180.6	4.8	44848	4	US-09-988-113-42	Sequence 42, Appli
	35	180.6	4.8	66933	4	US-09-544-398B-11	Sequence 11, Appli
	36	180.6	4.8	66933	4	US-09-543-771-11	Sequence 9, Appli
	37	180.6	4.8	72049	4	US-09-544-398B-9	Sequence 9, Appli
	38	180.6	4.8	72049	4	US-09-543-771-9	Sequence 7, Appli
	39	180.4	4.8	72604	3	US-09-268-992-7	Sequence 10, Appli
	40	180.4	4.8	72604	3	US-09-657-474-7	Sequence 7, Appli
	41	180.4	4.8	99500	3	US-09-798-096-10	Sequence 2252, Ap
	42	180	4.8	529	4	US-09-621-976-2252	Sequence 3, Appli
	43	179.6	4.8	66804	4	US-09-740-041-3	Sequence 10146, A
	44	179.6	4.8	368	4	US-09-621-976-10146	Sequence 6, Appli
	45	179.6	4.8	2240	4	US-09-814-351-6	

## ALIGNMENTS

RESULT 1  
US-09-584-568C-1  
; Sequence 1, Application US/09584568C  
; Patent No. 6500657  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria, Alexandra et al.  
; TITLE OF INVENTION: 33167, A NOVEL HUMAN HYDROLASE AND USES THEREFOR  
; FILE REFERENCE: MNI-140  
; CURRENT APPLICATION NUMBER: US/09/584,568C  
; CURRENT FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/193,954  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1332  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (47)..(991)  
US-09-584-568C-1

Query Match	5.0%	Score 187.8;	DB 4;	Length 1332;
Best Local Similarity	76.2%	Pred. No. 7.2e-29;		
Matches	231;	Conservative	0;	Mismatches 72;
				Indels 0;
				Gaps 0;
Qy	2306	AGATGTTAAGCTGTCATGCGAGTTCATGCCAGTGATCCGAGTACTTTGGAGCAAGAC	2365	
Db	1007	AGAAGATGCGCGGCGTGGTCTCATCTGTAATTCAGCACTTTGGAGGCTAAGGC	1066	
Qy	2366	AGGTGGATCTCTTTGAGCCAGGAGTTTGGCCAGACTGCACAACACAGTGAGACCTCGT	2425	
Db	1067	GGGAGGATGATTTGAGCCAGGAGTTGGAGACCAGCTGCCCAACATGTGAAACCTGT	1126	
Qy	2426	TTCTACAAATAATTAATAAATTTAGCCAGGTTGGTGTCACACCTGTAGTCCAGGTAC	2485	
Db	1127	CTCTACTAAATAATTAATAAATTTAGCCCTGGTGGTGTCACACCTGTATATCCAGGTAC	1186	
Qy	2486	TCAGGAGGCTGAGTGTAGTATGATTGTTGAGACTGGGAGGTTGAGGCTGAACCTGAGCCAG	2545	
Db	1187	TCGGAGGCTGAGCAGGAGAACTACTTGAACCTGGAGCAGAGGTTGCAATGAGCCGA	1246	
Qy	2546	GATCTTGGCCACCAATTCAGCTTTGGGCAACAGAGTGAGACCTGTCTCAAAAAAAAAA	2605	
Db	1247	GATCACCACCTACTACCTCCAGCTTAGGCAACAGAGCAAGACTCTGTCTCAAAAAAAAAA	1306	
Qy	2606	AAA 2608		
Db	1307	AAA 1309		

**This Page Blank (uspto)**

us-10-005-907-2.rge

Fri Nov 19 05:43:45 2004

Search completed: November 19, 2004, 02:51:07  
Job time : 3587 secs

Db 404 GGGAACTCTGCTGAAGACTATGAGATGTTCCCTGCAAGCTGAGAGACCCAGAGAG 463  
 Qy 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121  
 Db 464 TCCTGGGAGGAAGTCTGAGACTGAGTATTCACCTTCTACATATGCTTCTACAGACCCAGG 523  
 Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 524 CATGCCCGATCCCGAAGAGATGATGAACTTCTCATGCCTCAC 568

RESULT 14  
 LOCUS CQ733169 2505 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 19103 from Patent WO02068579.  
 ACCESSION CQ733169  
 VERSION CQ733169.1 GI:42315682  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 TITLE Kits, such as nucleic acid arrays, comprising a majority of  
 humanexons or transcripts, for detecting expression and other uses  
 thereof  
 JOURNAL Patent: WO 02068579-A 19103 06-SEP-2002;  
 PE Corporation (NY) (US)

FEATURES  
 source  
 Location/Qualifiers  
 1..2505  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.35e-07 Length: 2505  
 Score: 166.50 Matches: 43  
 Percent Similarity: 58.87% Conservative: 30  
 Best Local Similarity: 34.68% Mismatches: 46  
 Query Match: 23.29% Indels: 5  
 DB: 6 Gaps: 4

US-10-005-907-2 (1-135) x CQ733169 (1-2505)

Qy 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33  
 Db 2068 GAACATTTGGAGAGGCAATTCAG-----AAGAAAGAAATGTTAAGGAGGAAAAATA 2121  
 Qy 34 Thr-Thr-PheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerSerThr 53  
 Db 2122 CTCAATTTTGAAGAAGG---CAAGATTCCCAAAACGAAATGAAAGAAATGTCATCTACT 2178  
 Qy 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsn 73  
 Db 2179 CCATCCAGCAACATGTTGACCAAGACTTCTACAGAGAGCTGCTATACCTTCATCAAT 2238  
 Qy 74 His---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyrGluAsnIle 92  
 Db 2239 CATCGGGTCTCTGTCAAGGCCATCAGGCACTGCTGCTGAAGAGTACTATGAGAATGTT 2298  
 Qy 93 AspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlnTyrAlaLeu 112  
 Db 2299 CCTGCAAGCTGAGAGACCCAGAGAGTCTTGGGAGGAAGTCTGAGACTGAGTATTCACCT 2358  
 Qy 113 LeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAspTyrGluVal 131  
 Db 2359 CTACATATGCTTCTACAGACCCAGGCGATGCCCGATGCCGAGATGAATGAACATT 2418  
 Qy 132 ValPheProHis 135  
 Db 2419 CTCATGCCTCAC 2430

## RESULT 15

LOCUS AX740457 1643 bp DNA linear PAT 10-MAY-2003  
 DEFINITION Sequence 46 from Patent WO02020756.  
 ACCESSION AX740457  
 VERSION AX740457.1 GI:30523621  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Stuart, J., Lincoln, S.E., Altus, C.M., Dufour, G., Chalup, M.,  
 Hillman, J.L., Jones, A.L., Yu, J.Y., Wright, R.J., Gietzen, D., Liu, T.,  
 Yap, P., Dahl, C.R., Momiya, M.G., Bradley, D., Rohatgi, S.,  
 Harris, B., Roseberry, Ann, M., Gerstin, B.H., Peralta, C.H.,  
 David, M.H., Panzer, Scott, R., Flores, V., Daffo, A., Marwaha, R.,  
 Chen, A.J., Chang, S.C., Au, A.P. and Inman, R.R.  
 TITLE Secretory molecules  
 JOURNAL Patent: WO 02020756-A 46 14-MAR-2002;  
 Incyte Genomics, Inc. (US)

FEATURES  
 source  
 Location/Qualifiers  
 1..1643  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: LG:018739.2:2000SEP08"

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.22e-07 Length: 1643  
 Score: 159.00 Matches: 47  
 Percent Similarity: 50.64% Conservative: 32  
 Best Local Similarity: 30.13% Mismatches: 55  
 Query Match: 22.24% Indels: 22  
 DB: 6 Gaps: 6

US-10-005-907-2 (1-135) x AX740457 (1-1643)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19  
 Db 1096 ATGGGAAATTCCTCTGCTGAGAGAAAACGCGCGCAGCAGAACACTCAAGAGATGCCTGG 1155  
 Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 1156 AATGTGAGATGCAAAAGCCCAACAGAGAACATCCAGATGCTGGGATCCCATATCGCT 1215  
 Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
 Db 1216 GAAGGTGTTTCTGCTTCCATGGAAAAATACTCATTTTGAAGAGG---CAAGAT 1272  
 Qy 43 GlnAspLysLysSerGlnGluValSerSerThrSer---AsnGlnGluAsnGluAsnGly 61  
 Db 1273 TCCAAAACGAAATGAAAGAAATGCTATCTCCATCCATCCAGCAGACAATGTTGACCAG 1332  
 Qy 62 SerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 80  
 Db 1333 ACCTACTCAGAGAGCTGTGCTATACCCATCAATCATCGGGTCTCTGTGTCAAGGCCA 1392  
 Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
 Db 1393 TCAGGGAACCTCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGAGCTGAGAGACCC 1452  
 Qy 101 GlnPheArgGluArgSerGluThrGlnTyrAlaLeuLeuArgThrSerValSerArgPro 120  
 Db 1453 GAGTCCCTGGGAGGAAGTCTGAGACTGAGTATTCATCTTANATATGCTCTTCTACAGACC 1512  
 Qy 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 1513 AGGCATGCCCGATCCCGAGAGATGAATGATGAATTTTTCATGCCTCAC 1560

Medicine, Stanford University, 269 Campus Drive, CCSR building,  
Room 1100, Palo Alto, CA 94305-5151, USA

FEATURES  
source

Location/Qualifiers  
1..1654  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
173..709  
/note="germinal center-associated lymphoma; contains  
modified ITAM motif, similar to mouse M17"  
/codon\_start=1  
/product="HCAL"  
/protein\_id="RAO21701.1"  
/db\_xref="GI:27335683"  
/translation="MGNSLLRNRRQNTQMPNVRMQSPKQTSRCWDHHAEGCF  
CLPWKKILIFKRDQSONERMSSTPIQDNVDQYSELCTILNHRVLCTRPNGS  
AEEYVNPCKAERPRESLGGTETYSLLHMPSTDPHARSPEDEYELLMPHRISSHF  
LQQRPLMAPSETQFSHL"

CDS

ORIGIN

Alignment Scores:  
Pred. No.: 4.16e-08 Length: 1654  
Score: 169.50 Matches: 47  
Percent Similarity: 51.61% Conservative: 33  
Best Local Similarity: 30.32% Mismatches: 54  
Query Match: 23.71% Indels: 21  
DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x AF521911 (1-1654)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19  
DB 173 ATGGGAAATTCCTGCTGAGAGAAACACAGGGCGCAGACAACTCAAGATCGCTTG 232  
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
DB 233 AATGTGAGATGCAAGACCCCAACAGACATCCAGATCGTGGATCACCATATCGCT 292  
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
DB 293 GAAGGTGTTCTGCTGCTTCCATGGAATAATACTATTGTAAGAGAGG---CAAGAT 349  
QY 43 GlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysLys 62  
DB 350 TCCAAAACGAAATGAAGATGATGATCTTCTCCATCCAGCAGCAATGTGACGAGC 409  
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81  
DB 410 TACTCAGAGAGCTGTGCTATACCTCATCAATCATCGGTGTTCTGTACAGGCCATCA 469  
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 101  
DB 470 GGGAACTGCTGCAAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCCAGAG 529  
QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 121  
DB 530 TCCTTGGGAGAACTGAGACTGAGTATCTACTTCTCATATGCTTCTACAGCCCGAG 589  
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
DB 590 CATGCCCGATCCCAAGAGATGATATGAACTTCTCATGCTCTCAC 634

RESULT 13

AY212246  
LOCUS 3270 bp mRNA linear PRI 14-OCT-2003  
DEFINITION Homo sapiens germinal center B-cell expressed transcript 2 (GCET2)  
mRNA, complete cds.  
ACCESSION AY212246  
VERSION AY212246.1 GI:27948576  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS

1 (bases 1 to 3270)  
Pan, Z., Shen, Y., Du, C., Zhou, G., Rosenwald, A., Staudt, L. M.,  
Greiner, T. C., McKeithan, T. W. and Chan, W. C.

TITLE

Two newly characterized germinal center B-cell-associated genes,  
GCET1 and GCET2, have differential expression in normal and  
neoplastic B cells  
Am. J. Pathol. 163 (1), 135-144 (2003)

JOURNAL

22702315

MEDLINE

12819018

PUBMED

2 (bases 1 to 3270)  
Pan, Z., Shen, Y., Du, C., Zhou, G., McKeithan, T., Rosenwald, A.,  
Staudt, L. and Chan, W. C.

AUTHORS

Direct Submission

TITLE

Submitted (07-JAN-2003) Pathology, University of Nebraska Medical  
Center, #42, Omaha, NE 68198, USA

JOURNAL

Location/Qualifiers

FEATURES

source

1..3270  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3q13.13"  
/cell\_line="DHL16"  
1..3270  
/gene="GCET2"  
/note="GCAT2"  
107..643  
/gene="GCET2"  
/note="similar to Mus musculus germinal center transcript  
M17; similar to MGC40441 sequence in GenBank Accession  
Number BC030506"  
/codon\_start=1  
/product="germinal center B-cell expressed transcript 2"  
/protein\_id="AA022147.1"  
/db\_xref="GI:27948577"  
/translation="MGNSLLRNRRQNTQMPNVRMQSPKQTSRCWDHHAEGCF  
CLPWKKILIFKRDQSONERMSSTPIQDNVDQYSELCTILNHRVLCTRPNGS  
AEEYVNPCKAERPRESLGGTETYSLLHMPSTDPHARSPEDEYELLMPHRISSHF  
LQQRPLMAPSETQFSHL"

gene

CDS

ORIGIN

Alignment Scores:  
Pred. No.: 8.77e-08 Length: 3270  
Score: 169.50 Matches: 47  
Percent Similarity: 51.61% Conservative: 33  
Best Local Similarity: 30.32% Mismatches: 54  
Query Match: 23.71% Indels: 21  
DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x AY212246 (1-3270)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19  
DB 107 ATGGGAAATTCCTGCTGAGAGAAACACAGGGCGCAGACAACTCAAGATCGCTTG 166  
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
DB 167 AACGTGAGATGCAAGACCCCAACAGACATCCAGATCGTGGATCACCATATCGCT 226  
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
DB 227 GAAGGTGTTCTGCTGCTTCCATGGAATAATACTATTGTAAGAGAGG---CAAGAT 283  
QY 43 GlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysLys 62  
DB 284 TCCAAAACGAAATGAAGATGATGATCTTCTCCATCCAGCAGCAATGTGACGAGC 343  
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81  
DB 344 TACTCAGAGAGCTGTGCTATACCTCATCAATCATCGGTGTTCTGTACAGGCCATCA 403  
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 101



Klausner, R.D., Collins, F.S., Wagner, L., Shennmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.F., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tohyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S., McEwan, P.J., McKernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., Worsley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Kryzhanovsk, M.I., Skalska, U., Smalusz, D.E., Sutterich, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1532)

Strausberg, R.

Direct Submission

Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

Gaithersburg, Maryland

Web site: <http://www.nisic.nih.gov/>

Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, O.L., Masiello, C., Masker, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Skarlip, S., Thomas, P.J., Touchman, J.W., Young, A.C., Vogt, J.H., Walker, M.A., Wetherby, K.D., Wiggins, L., Tsung, A., Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 64 Row: a Column: 15  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 227495336.

FEATURES  
source

```

Location/Qualifiers
1. .1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="MGC:40441 IMAGE:4385178"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH MGC 85"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1. .1532
/gene="GCET2"
/note="synonyms: MGC40441, HGAL, GCAT2"
/db_xref="LocusID:257144"
/db_xref="MIM:607792"
54. .590
/gene="GCET2"
/codon_start=1
/product="germinal center expressed transcript 2"
/protein_id="AAH30506.1"

```

```

/db_xref="GI:20987806"
/db_xref="LocustID:257144"
/db_xref="MIM:607792"
/translation="MGNSLLRENRRQNTQEMPNVWVMSQPKRSCVDHHAIEGCF
CLPWKILIFKRQSQNENRMSSTPIQNDVDQTYSELCYTLNHRVLCRTPSGNS
ABEYENVPCKAEPRRESLCGGTETYSLLHMPSTDPRHARSPEDRYELLMPHRISSHF
LQQPRPLMAPSTQFSHL"

```

## ORIGIN

Alignment Scores:	3.82e-08	Length:	1532
Pred. No.:	169.50	Matches:	47
Score:	51.61%	Conservative:	33
Percent Similarity:	30.32%	Mismatches:	54
Best Local Similarity:	23.71%	Indels:	21
Query Match:	9	Gaps:	5
DB:			

US-10-005-907-2 (1-135) x BC030506 (1-1532)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19

20 ---LysLysGlyAsnProAspGluGluargLysArg----- 30  
 QY

114 AATGTGAGAAATGCAAGAGCCCAACAGAGAACATTCAGATGCTGGGATCACCATATCGCT 173

Db 174 GAAGGGTGTTCGCGCTTCCATGGAAAAAATACTCATTTTGGAAAGAGG---CAAGAT 230

43	GlnAsp	Phys	Lyse	Ser	Gln	Glu	Val	Ser	Ser	Thr	Asn	Gln	Glu	Asn	Glu	Asn	Gly	Ser	62
Db	231	TCCCAAA	CGAAA	ATGAA	AGATG	TCAT	TACT	TCCAT	TCCAG	GACAT	TGTT	CAAC	CAGAC	CC					

[illegible]

**OY**

82 LeuSerSerAsnAspGlyTyrGluAsnIleaspSerLeuThrArgLysValArgGln 101  
:::| | | | | . . . . .  
.....

DB	351	GGGAAC	TCTG	TGAAG	AGTACT	ATGAGA	ATGTTCC	CTCG	AAGCT	GAGAC	CCGAG	GAG	410									
QY	102	Phe	Arg	Glu	Arg	Ser	Glu	Thr	Glu	Tyr	Ala	Leu	Leu	Arg	Thr	Ser	Val	Ser	Arg	Pro	Cys	121

Dbb 411 TCCTGGGAGAACTGAGACTGAGTATTCACTTCTACATATGCCTTCTACAGACCCCAGG 470

dbb 471 CATGCCCGATCCCAGAGATGAATATGAACCTTCTCATGCCCTCAC 515

	AFS21911	LOCUS	1654 bp	mRNA	liver	BBT 14 JAN 2002
RESULT 12	AFS21911					

DEFINITION	Homo sapiens HGAL mRNA, complete cds.
ACCESSION	AF521911
VERSION	AF521911.1 GI:27733682

ASIWKORUS  
SOURCE  
Homo sapiens (human)  
Homo sapiens  
ORGANISM

REFERENCE  
 ALLEN, G. D. 1961. The  
 Loricata of the  
 1 (bases 1 to 1654)  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 Craniata; Vertebrata; Euteleostomi;  
 Mollusca; Eutherozoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

**TITLE** HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma  
**JOURNAL** Blood 101 (2) 432-440 (2003)

MEDLINE	22397484
PUBMED	12509382
REFERENCE	2 (bases 1 to 1654)

**AUTHORS** Lossos, I.S., Alizadeh, A.A. and Levy, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (16-JUN-2002) Division of Oncology, Department of

```

/misc_feature /note="assembly_name:Contig2"
64946. .67606
/note="assembly_name:Contig3"
67707. .77674
/note="assembly_name:Contig4"
77775. .89256
/note="assembly_name:Contig5"
89357. .100400
/note="assembly_name:Contig6"
100501. .111915
/note="assembly_name:Contig7"
112016. .125713
/note="assembly_name:Contig8"
125814. .141268
/note="assembly_name:Contig9
clone_end:SP6
vector_side:left"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 2.44e-33 Length: 141268
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
DB: 2 Gaps: 0

```

US-10-005-907-2 (1-135) x AC074365 (1-141268)

```

QY 56 GlnGluAenGluAenGlySerGlySerGluGluValCysTyrThrValIleAenHisIle 75
DB 5700 CAGGAAACGAGAAATGGCAGTGTCTTGAAGAGTGTCTACACTGCTATTATACATC 5641

QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAenIleAspSerLeu 95
DB 5640 CCCCATCAGAGATCCCTCCCTGAGCTCCATGATGGCTATGAGAACATTCACCTCCCTC 5581

QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
DB 5580 ACAAGAAAGTGTAGACAGATTTAGAGAAAGTTCAGACAGAAATATGCCCTTCTTAGACT 5521

QY 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 5520 TCTGTTAGTAGGCTTGTCTCTCCATCCATGAGCATGATTAAGAGTTGTGTTTCCACAC 5461

```

## RESULT 10

```

AL606804 185467 bp DNA linear PRI 23-OCT-2002
LOCUS Human DNA sequence from clone RP11-978115 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL606804
VERSION AL606804.11 GI:24366459
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Almeida, J.
Direct Submission
Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
Cloned from: Cloned from: Cloned from: Cloned from: Cloned from: Cloned from:
On Oct 24, 2002 this sequence version replaced gi:18121511.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

```

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep). This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>. RP11-978115 is from the library RPCI-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

FEATURES  
source

```

----- Location/Qualifiers
1..185467
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-978115"
/clone_lib="RPCI-11.4"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 3.29e-33 Length: 185467
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
DB: 9 Gaps: 0

```

US-10-005-907-2 (1-135) x AL606804 (1-185467)

```

QY 56 GlnGluAenGluAenGlySerGlySerGluGluValCysTyrThrValIleAenHisIle 75
DB 132080 CAGGAAACGAGAAATGGCAGTGTCTTGAAGAGTGTCTACACTGCTATTATACATC 132139

QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAenIleAspSerLeu 95
DB 132140 CCCCATCAGAGATCCCTCCCTGAGCTCCATGATGGCTATGAGAACATTCACCTC 132199

QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
DB 132200 ACAAGAAAGTGTAGACAGATTTAGAGAAAGTTCAGACAGAAATATGCCCTTCTTAGACT 132259

QY 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 132260 TCTGTTAGTAGGCTTGTCTCTCCATCCATGAGCATGATTAAGAGTTGTGTTTCCACAC 132319

```

## RESULT 11

```

BC030506 1532 bp mRNA linear PRI 29-JUN-2004
LOCUS BC030506
DEFINITION Homo sapiens germinal center expressed transcript 2, mRNA (cdna
Clone MGC:40441 IMAGE:4385178), complete cds.
ACCESSION BC030506
VERSION BC030506.1 GI:20987805
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1532)
REFERENCE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
AUTHORS

```

Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.

#### FEATURES

Location/Qualifiers  
1..4136  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRACE3002390"  
/issue\_type="cerebellum"  
/clone\_lib="BRACE3"  
/note="cloning vector: pME18SFL3"  
136..537  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAC85873.1"  
/db\_xref="GI:34530325"  
/translation="MWLEDSNSTGCGARNVCVCLVFLSVLGCQSRKQGMRTQQAG  
RWLPAGREASSETNPEGLTLDPSFLPAAAGWQHEPVRVNCVLPHPAFSSWCFCQ  
NFKWSETQLLNRTVTKRWEIISCNSGRK"

#### ORIGIN

Alignment Scores:  
Pred. No.: 7.16e-51 Length: 4136  
Score: 581.00 Matches: 114  
Percent Similarity: 85.19% Conservative: 1  
Best Local Similarity: 84.44% Mismatches: 0  
Query Match: 81.26% Indels: 20  
DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK124520 (1-4136)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20  
Db 498 ATGGGAAATTATCTCTCGGAAACTC----- 524  
QY 21 LysGlyAsnProaspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
Db 525 -----AGGCAGGAAATGACTACATTTGAAAGAAACTT 557  
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60  
Db 558 CAAGATCAAGATGAAGAAAGCCAGAGAGTTTCACCTCTCTATATCAGGAAACGAGAA 617  
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80  
Db 618 GGCAGTGGTTCTGAAGAAGTGTGTACACTGTCTAATACATCCCCCATCAAAATCC 677  
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
Db 678 TCCTGAGCTCCATGATGATGCTATGAGAACTTGAACATTCCTCCATCAAGGAAAGTGA 737  
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
Db 738 CAGTTTAGAGAAAGGTCAGAGACAGAAATAGCCCTCTTAGGACTCTCTGTAGGCCT 797  
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135  
Db 798 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 842

#### RESULT 9

AC074365/c  
LOCUS AC074365 141268 bp DNA linear HTG 23-SEP-2000  
DEFINITION Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,  
10 unordered pieces.  
AC074365  
ACCESSION AC074365  
VERSION AC074365.5 GI:10280935  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

#### REFERENCE

AUTHORS Waterston,R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 141268)

AUTHORS Waterston,R.H.

TITLE Direct Submission

JOURNAL Submitted (29-JUL-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

COMMENT On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H NH0115C04  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Chemistry: Dye-terminator; plasmid: 0%  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 136274 bases at least Q40  
Consensus quality: 138356 bases at least Q30  
Consensus quality: 139182 bases at least Q20  
Insert size: 142000; agarose-fp  
Insert size: 140368; sum-of-contigs  
Quality coverage: 4.64 in Q20 bases; agarose-fp  
Quality coverage: 4.81 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved

\* 1 31377: contig of 31377 bp in length  
\* 31378 31477: gap of unknown length  
\* 31478 63011: contig of 31534 bp in length  
\* 63012 63111: gap of unknown length  
\* 63112 64845: contig of 1734 bp in length  
\* 64846 64945: gap of unknown length  
\* 64946 67066: contig of 2661 bp in length  
\* 67067 67707: gap of unknown length  
\* 67707 77674: contig of 9968 bp in length  
\* 77675 89256: contig of 11482 bp in length  
\* 89257 89357: gap of unknown length  
\* 89357 100400: contig of 11044 bp in length  
\* 100401 100500: gap of unknown length  
\* 100501 111915: contig of 11415 bp in length  
\* 111916 112015: gap of unknown length  
\* 112016 125713: contig of 13698 bp in length  
\* 125714 125813: gap of unknown length  
\* 125814 141268: contig of 15455 bp in length.

#### FEATURES

source  
1..141268  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-115C4"  
1..31377  
/note="assembly\_name:Contig10"  
31478..63011  
/note="assembly\_name:Contig11"  
63112..64845  
misc\_feature  
misc\_feature  
misc\_feature

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 141268)  
Waterston,R.H.

HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: Reverse Proteomics Research Institute, HRI and RAB.

# FEATURES

source  
Location/Qualifiers  
1..4032  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRACE3024537"  
/tissue\_type="cerebellum"  
/clone\_lib="BRACE3"  
/note="Cloning vector: pME18SFL3"

## ORIGIN

Alignment Scores:  
Pred. No.: 3..37e-51 Length: 4032  
Score: 584.00 Matches: 115  
Percent Similarity: 85.19% Conservatives: 0  
Best Local Similarity: 85.19% Mismatches: 0  
Query Match: 81.68% Indels: 20  
DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK126682 (1-4032)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20  
DB 377 ATGGGAATTAATCTCTCTCGAAATCTC----- 403

QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
DB 404 -----AGGCAGAAATGACTACATTTGAAAGAAACTT 436

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60  
DB 437 CAAGATCAGATAAGAAAGCCAGAGTTTCTATCCACTTCTTAATCAGGAAACGAGAAT 496

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80  
DB 497 GGCAGTGGTTCTGAAGAAGTGTCTACACTGTCTAATTAATCAGATCCCATCAGATCC 556

QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
DB 557 TCCTGAGCTCCAAATGATGCTATGAGAACATTGACTCCCTCACAGGAAAGTGAGA 616

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
DB 617 CAGTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGAGCCT 676

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135  
DB 677 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 721

## RESULT 7

CQ841482  
LOCUS CQ841482 4136 bp DNA linear PAT 02-AUG-2004  
DEFINITION Sequence 129 from Patent EP1440981.  
ACCESSION CQ841482  
VERSION CQ841482.1 GI:50893269  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1  
Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.  
Full-length human cdna  
Patent: EP 1440981-A 129 28-JUL-2004;  
Research Association for Biotechnology (JP)  
Location/Qualifiers  
1..4136  
/organism="Homo sapiens"

REFERENCE 2  
Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.  
Full-length human cdna  
Patent: EP 1440981-A 129 28-JUL-2004;  
Research Association for Biotechnology (JP)  
Location/Qualifiers  
1..4136  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
Pred. No.: 7..16e-51 Length: 4136  
Score: 581.00 Matches: 114  
Percent Similarity: 85.19% Conservatives: 1  
Best Local Similarity: 84.44% Mismatches: 0  
Query Match: 81.26% Indels: 20  
DB: 6 Gaps: 1

US-10-005-907-2 (1-135) x CQ841482 (1-4136)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20  
DB 498 ATGGGAATTAATCTCTCTCGAAATCTC----- 524

QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
DB 525 -----AGGCAGGAAATGACTACATTTGAAAGAAACTT 557

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60  
DB 558 CAAGATCAGATAAGAAAGCCAGAGTTTCTATCCACTTCTTAATCAGGAAACGAGAAT 617

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80  
DB 618 GGCAGTGGTTCTGAAGAAGTGTCTACACTGTCTAATTAATCAGATCCCATCAGAAATCC 677

QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
DB 678 TCCTGAGCTCCAAATGATGCTATGAGAACATTGACTCCCTCACAGGAAAGTGAGA 737

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
DB 738 CAGTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGAGCCT 797

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135  
DB 798 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 842

## RESULT 8

AK124520  
LOCUS AK124520 4136 bp mRNA linear PRI 09-SEP-2003  
DEFINITION Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.  
ACCESSION AK124520  
VERSION AK124520.1 GI:34530324  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1  
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, S., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kikuchi, H., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 4136)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
(E-mail: genomics@hri.co.jp, Tel. 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing:

```

Db      299 TCCCTGAGCTCCAAATGATGCGCTATGAGAACATTGACTCCCTCACAGGAAGTGAGA 358
Qy      101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db      359 CAGTTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTTCTTAGACTTCTGTAGTAGGCT 418
Qy      121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db      419 TGTTCCTGCACCCATGACATGATATGAAAGTTGTGTTCACAC 463

RESULT 5
AB060908
LOCUS      AB060908
DEFINITION Macaca fascicularis brain cDNA clone:Qtra-14007, full insert
sequence.
ACCESSION AB060908
VERSION    AB060908.1 GI:13874585
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE     Macaca fascicularis (crab-eating macaque)
ORGANISM   Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
            Cercopithecinae; Macaca.
REFERENCE 1
AUTHORS    Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirata,M., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
            Prediction of unidentified human genes on the basis of sequence
            similarity to novel cDNAs from cynomolgus monkey brain
            (ex) Genome Biol. 3, research0006.1-0006.5 (2001)
            2 (bases 1 to 1924)
            Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
            Direct Submission
            Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashienih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
            Lab host: TOP10
            Vector: pME18S-FL3 (Acc.No. AB009864)
            R. Site: DraIII (CAGCTGTGTG)
            R. Site: DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGGCTGCTTTTCTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR. The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert. Libraries
            were constructed by oligo-capping method
            (Sugano et al., Institute of Medical Science, University of
            Tokyo).
            Custom primer used for sequencing
            (5' end primer [CTTCTGCTCTAAAGCTGCG];
            3' end primer [CGACCTCGACCTGAGCACA] ).
            Location/Qualifiers
            1..1924
            /organism="Macaca fascicularis"
            /mol_type="mRNA"
            /db_xref="taxon:9541"
            /clone="Qtra-14007"
            /sex="male"
            /tissue type="temporal lobe right"
            /clone_lib="macaque brain cDNA library Qtra"
            /dev_stage="adult"
            248..559
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="BAB46903.1"
            /db_xref="GI:13874586"
            /translation="WTFPERKLDQDKSKVSSISNQENNGSGSEEVYTVINHP
            HRKSSLSNDGVDYENIDSLTRKREFRSEYFALLRTSVSRPYSCTHEHDYEVLP
            H"

FEATURES
source
CDs
ORIGIN

```

## Alignment Scores:

Pred. No.: 8.44e-57 Length: 1924  
 Score: 634.00 Matches: 120  
 Percent Similarity: 96.12% Conservative: 4  
 Best Local Similarity: 93.02% Mismatches: 5  
 Query Match: 88.67% Indels: 0  
 DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AB060908 (1-1924)

```

Qy      7  ArgLysLeuSerCysLeuGluAsnGlnLysLysProLysLysGlyAsnProAspGlu 26
Db      170 AGGACATGGAGTTGCCCTGGGAGAGAAATCAAAAGAAAGCCAGGAAGAAACCCAGATGAG 229
Qy      27  GluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLys 46
Db      230 GAAAGAAAACGGCAGGAATGACTACATTTTGAAGAAAACCTTCAAGATCAAGATAAGAAA 289
Qy      47  SerGlnGluValSerSerThrSerAsnGlnGluAsnGlnGlySerGlySerGluGlu 66
Db      290 AGCAAGAAGTTTCATCCATTCTTAATCAGGAAAACGAGAATGCGAGTGGTCTCTGAAGAA 349
Qy      67  ValCysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsp 86
Db      350 GTGTGCTACACTGTCAATTAATCATCCCCATCGGAGGCTTTCCTCGAGCTCCATGAT 409
Qy      87  AspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106
Db      410 GATGGCTATGAGAACATTGACTCCCTCACAGGAAGTGAGAAATTTAGAAAGTCA 469
Qy      107 GluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126
Db      470 GAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCTTATCTTCGACCCATGAG 529
Qy      127 HisAspTyrGluValValPheProHis 135
Db      530 CATGATTACGAAGTTGTGTTCACAC 556

```

## RESULT 6

AK126682 4032 bp mRNA linear PRI 19-FEB-2004  
 Homo sapiens cDNA FL44728 fis, clone BRACE3024537.  
 AK126682  
 DEFINITION oligo capping; fis (full insert sequence).  
 ACCESSION AK126682.1 GI:34533254  
 VERSION  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,  
 Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,  
 Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,  
 Sugiyama,T., Irie,K., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,  
 Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,  
 Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,  
 Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,  
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
 Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 4032)  
 Isogai,T. and Yamamoto,J.  
 Direct Submission  
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
 (E-Mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

```

Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 385 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 429

RESULT 3
CQ842314
LOCUS CQ842314 Homo sapiens (human) linear PAT 02-AUG-2004
DEFINITION Sequence 961 from Patent EP1440981.
ACCESSION CQ842314
VERSION CQ842314.1 GI:50894101
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yanamoto,J., Isono,Y., Nagai,K. and Irie,R.
TITLE Full-length human cDNA
JOURNAL Patent: EP 1440981-A 961 28-JUL-2004;
RESEARCH Association for Biotechnology (JP)
FEATURES
source
1..1709
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6,12e-65 Length: 1709
Score: 711.00 Matches: 134
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.26% Mismatches: 0
Query Match: 99.44% Indels: 0
DB: 6 Gaps: 0
US-10-005-907-2 (1-135) x CQ842314 (1-1709)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
Db 59 ATGGGAATTAATCTCTCGGAAACTCAGTTCCTGGGAGAGAAATCAAAAGAACGCCAAG 118
Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 119 AAAGGAAACCCAGATGAGGAAAGAAACCGCAGAGAAATGACTACATTTGAAGAAACCTT 178
Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
Db 179 CAAGATCGAGATAAGAAAGCCCAAGATTTTCATCCACTTCTAATCAGGAAACGAGAAT 238
Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 239 GGCAGTGGTTCGAGAGAGTGCTACACTGTCTAATACATCCATCCCATCAGAGATCC 298
Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 419 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 463

RESULT 4
AK123798
LOCUS AK123798 Homo sapiens cDNA FLJ41804 f1s, clone NOVAR2000710.
DEFINITION
ACCESSION AK123798
VERSION AK123798.1 GI:34529425
KEYWORDS oligo capping; f1s (full insert sequence).

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T.,
Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R.,
Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H.,
Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E.,
Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S.,
Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T.,
Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1709)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..1709
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NOVAR2000710"
/tissue_type="ovary"
/clone_lib="NOVAR2"
/dev_stage="adult"
/note="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 6,12e-65 Length: 1709
Score: 711.00 Matches: 134
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.26% Mismatches: 0
Query Match: 99.44% Indels: 0
DB: 6 Gaps: 0
US-10-005-907-2 (1-135) x AK123798 (1-1709)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
Db 59 ATGGGAATTAATCTCTCGGAAACTCAGTTCCTGGGAGAGAAATCAAAAGAACGCCAAG 118
Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 119 AAAGGAAACCCAGATGAGGAAAGAAACCGCAGAGAAATGACTACATTTGAAGAAACCTT 178
Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
Db 179 CAAGATCGAGATAAGAAAGCCCAAGATTTTCATCCACTTCTAATCAGGAAACGAGAAT 238
Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 239 GGCAGTGGTTCGAGAGAGTGCTACACTGTCTAATACATCCATCCCATCAGAGATCC 298
Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100

```

```

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2648)
Strausberg,R.
Direct Submission
Submitted (19-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systembiology.org
contact: amadan@systembiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 32 Row: b Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21687148.
FEATURES             source
    1..2648
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4109064"
        /tissue_type="Bone marrow, acute myelogenous leukemia"
        /clone_lib="NIH_MGC_55"
        /lab_host="DH10B"
        /note="Vector: pDNR-LIB"
ORIGIN
Alignment Scores:
Pred. No.:          3,76e-65          Length:          2648
Score:              715.00           Matches:          135
Percent Similarity: 100.00%          Conservative:      0
Best Local Similarity: 100.00%        Mismatches:       0
Query Match:        100.00%          Indels:           0
DB:                 9                Gaps:             0
US-10-005-907-2 (1-135) x BC024174 (1-2648)
QY      1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysProLys 20
Db      48 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAG 107
QY      21 LysGlyAsnProaspGluGluAtrGlyArgGluGlnMetThrPheGluArgLysLeu 40
Db      108 AAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACTT 167
QY      41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAenGluAen 60
Db      168 CAAGATCAAGATAAGAAAGCCCAAGAGTTTTCATCCACTTCTAATCAGGAAACGAGAAAT 227
QY      61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db      228 GGCAGTGGTTCGAGGAAGTGTGTACACTGTCTAATTAATCACATCCCCATCAGAGATCC 287
QY      81 SerLeuSerSerAsnAspAspGlyTyrGluAenIleAspSerLeuThrArgLysValArg 100
Db      288 TCCTGAGCTCCCAATGATGAGCTATGAGAAATTCATTGACTCCCTCACAAGGAAAGTGAGA 347
QY      101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120

```

---

```

121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
408 TGTTCCTGCACCCATGAGCATGATTGAGAGTTGTGTTCACAC 452
RESULT 2
AX505122
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
    source
        1..3762
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            25..432
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD48808.1"
            /db_xref="GI:23386430"
            /translation="MGNYLLRLKSLCLGENOKKPKGNPDERKEQEMTTTPEKLDQD
KKSOEVSSTSNQENENGSGSEEVYVINIHPQRSLSNSDDGYENIDSLTRKVRQF
RERSEYALLRTSVSRPCSTHEHDYEVVFP"
ORIGIN
Alignment Scores:
Pred. No.:          5,52e-65          Length:          3762
Score:              715.00           Matches:          135
Percent Similarity: 100.00%          Conservative:      0
Best Local Similarity: 100.00%        Mismatches:       0
Query Match:        100.00%          Indels:           0
DB:                 6                Gaps:             0
US-10-005-907-2 (1-135) x AX505122 (1-3762)
QY      1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysProLys 20
Db      25 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAG 84
QY      21 LysGlyAsnProaspGluGluArgLysArgGlnGluMetThrPheGluArgLysLeu 40
Db      85 AAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACTT 144
QY      41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAenGluAen 60
Db      145 CAAGATCAAGATAAGAAAGCCCAAGAGTTTTCATCCACTTCTAATCAGGAAACGAGAAAT 204
QY      61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db      205 GGCAGTGGTTCGAGGAAGTGTGTACACTGTCTAATTAATCACATCCCCATCAGAGATCC 264
QY      81 SerLeuSerSerAsnAspAspGlyTyrGluAenIleAspSerLeuThrArgLysValArg 100
Db      265 TCCTGAGCTCCCAATGATGAGCTATGAGAAATTCATTGACTCCCTCACAAGGAAAGTGAGA 324
QY      101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db      325 CAGTTTAGAGAAAGGTGAGAGACAGAGATATGCCCTTCTTGGACTTCTCTGTAGTAGGCTT 384

```



## ALIGNMENTS

Alignment Scores:

Pred. No.:	1.11e-07	Length:	876
Score:	156.50	Matches:	37
Percent Similarity:	60.75%	Conservative:	28
Best Local Similarity:	34.58%	Mismatches:	39
Query Match:	21.89%	Indels:	3
DB:	5	Gaps:	3

US-10-005-907-2 (1-135) x BQ214406 (1-876)

QY	31	GlnGluMetThrThrPheGluArgLeuGlnAspGlnAspLysSerGlnGluVal	50
Db	384	AAAAAAATACTCATTTTTTAAAGAGG--CAAGATTCCCAAAACGAAATGAAGAAGT	440
QY	51	SerSerThrSerAsnGlnGluAsnGlySerGlySerGluGluValCysTyrThr	70
Db	441	TCATCTACTCCCATCCAGGACAAATGTGACAGACCTACTTCAGAGAGCTGTGCTATACC	500
QY	71	ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspSgLyTyr	89
Db	501	CTCATCAATCATCGGTTCTCTGTACAAAGGCCATCAGGAACTCTCTGAAGAGTACTAT	560
QY	90	GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu	109
Db	561	GAGAATGTTCCCTCCAAAGCTGAGAGACCAGAGAGTCTTGGGAGGAATCAGAGCTGAG	620
QY	110	TyrAlaIleuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp	128
Db	621	TATTCACTTCTACATATGCGCTTCTACAGACCCCGAGGCATGCCCGATCCCCAGAGATGAA	680
QY	129	TyrGluValValPheProHis	135
Db	681	TATGAATTTCTCATGCGCTCAC	701

Search completed: November 19, 2004, 03:36:44  
Job time : 2710 secs

Percent Similarity: 60.75% Conservative: 28  
 Best Local Similarity: 34.58% Mismatches: 39  
 Query Match: 21.89% Indels: 3  
 DB: 2 Gaps: 3

US-10-005-907-2 (1-135) x AW963366 (1-683)

Qy 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50  
 Db 57 AAAAAAATACCTATTTTGAAGAAGAGG---CAAGATTCCCAAAACGAAAAATGAAGAATG 113  
 Qy 51 SerSerThrSerAsnGlnGluAsnGlnGluAsnGlnGluSerGlySerGluGluValCysThr 70  
 Db 114 TCATCTACTCCCATCCAGGCAATGTTGACCAAGACCTACTCAGAGGAGCTGTGCTATACC 173  
 Qy 71 VallieAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyr 89  
 Db 174 CTCATCAATCATCGGGTTCCTGTACAAAGCCATCAGGAGACTCTGCTGAAGAGTACTAT 233  
 Qy 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109  
 Db 234 GAGAAATGTTCCCTGCAAGCTGAAAGCCAGAGAGTCTTGGGAGGAACTCAGACTGAG 293  
 Qy 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128  
 Db 294 TATTCACCTTCTACATATGCTTCTACAGACCCAGGATGCCCGATCCCGAAGATGAA 353  
 Qy 129 TyrGluValValPheProHis 135  
 Db 354 TATGAACCTTCTATGCTCTCAC 374

RESULT 15

BQ214406

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 876)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaaps-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

plate: L14M13353 row: m column: 14

High quality sequence stop: 678.

Location/Qualifiers

1. .876

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6070069"

/tissue\_type="embryonal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores: 5.51e-08 Length: 508  
 Pred. No.: 156.50 Matches: 37  
 Score: 156.50 Conservative: 28  
 Percent Similarity: 60.75% Mismatches: 39  
 Best Local Similarity: 34.58% Indels: 3  
 Query Match: 21.89% Gaps: 3  
 DB: 9

US-10-005-907-2 (1-135) x AY404586 (1-508)

Qy 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50  
 Db 116 AAAAAAATACCTATTTTGAAGAAGAGG---CAAGATTCCCAAAACGAAAAATGAAGAATG 172  
 Qy 51 SerSerThrSerAsnGlnGluAsnGlnGluAsnGlnGluSerGlySerGluGluValCysThr 70  
 Db 173 TCATCTACTCCCATCCAGGCAATGTTGACCAAGACCTACTCAGAGGAGCTGTGCTATACC 232  
 Qy 71 VallieAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyr 89  
 Db 233 CTCATCAATCATCGGGTTCCTGTACAAAGCCATCAGGAGACTCTGCTGAAGAGTACTAT 292  
 Qy 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109  
 Db 293 GAGAAATGTTCCCTGCAAGCTGAGAGCCAGGAGTCTTGGGAGGAACTCAGACTGAG 352  
 Qy 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128  
 Db 353 TATTCACCTTCTACATATGCTTCTACAGACCCAGGATGCCCGATCCCGAAGATGAA 412  
 Qy 129 TyrGluValValPheProHis 135  
 Db 413 TATGAACCTTCTATGCTCTCAC 433

RESULT 14

AW963366

LOCUS

DEFINITION

EST375439 MAGG resequences, MAGH Homo sapiens cDNA, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 683)

Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,

Holt, J. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and

Quackenbush, J.

Assessment of gene expression patterns in a model of colon tumor

metastasis using a 19,200 element cDNA microarray

Unpublished (2000)

Contact: John Quackenbush

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 3528

Fax: 301 838 0208

Email: johnq@tigr.org

Plate: 187

Seq primer: Reverse.

Location/Qualifiers

1. .683

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="MAGG resequences, MAGH"

/note="vector: pBluescriptSKM"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 8.08e-08 Length: 683

Score: 156.50 Matches: 37

QY 19 Pro-----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 125 CTTTGAATGTCAGATGCAAGGCCCAACAGAGAACATCCAGATGCTGGATCACCAT 184  
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeu 40  
 Db 185 ATCGCTGAAGGCTGTTCTGCTCCATGTAAGAAATACTCATTTTGAAGAGG--- 241  
 QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60  
 Db 242 CAAGATCCCAAAACGAAAGATGCAATCTACTCCATCCAGGCAATGTCAC 301  
 QY 61 GlySerGlySerGluGluValCysThrThrValIleAsnHis---IleProHisGlnArg 79  
 Db 302 CAGACCTACTCAGAGGAGCTGTGTATACCTCATCATCTCGGTTCTCTGTACAAGG 361  
 QY 80 SerSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99  
 Db 362 CCATCAGGGAATCTCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGCTGAGAGACC 421  
 QY 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119  
 Db 422 AGAGAGCTCTGGAGGAGTACTGAGTATGATTCATCTTACATATGCTTCTACAGAC 481  
 QY 120 ProCysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 482 CCCAGGATGCCGATCCCAAGAGATGAATATGAACTTCTCATGCTCTCAC 532

RESULT 12  
 BM455407  
 LOCUS  
 DEFINITION  
 AGENCOURT 6407334 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:5500391  
 5', mRNA sequence.  
 BM455407  
 BM455407.1 GI:18504436  
 EST.  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@email.nih.gov  
 Tissue Procurement: Lou Staudt  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Cloning Distribution: Agencourt Bioscience Corporation  
 Cloning Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLNL2135 row: d column: 24  
 High quality sequence start: 19  
 High quality sequence stop: 760.  
 Location/Qualifiers  
 1..1085  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5500391"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_85"  
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.867 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 6.55e-08 Length: 1085  
 Score: 159.50 Matches: 37  
 Percent Similarity: 60.75% Conservative: 28  
 Best Local Similarity: 34.58% Mismatches: 39  
 Query Match: 22.31% Indels: 3  
 DB: 4 Gaps: 3  
 US-10-005-907-2 (1-135) x BM455407 (1-1085)  
 QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50  
 Db 439 AAAAAATATCATTTTGAAGAGG---CAAGATCCCAAAACGAAATGAAGATG 495  
 QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGluGluValCysThrThr 70  
 Db 496 TCATCTACTCCATCCAGGACAATGTTGACCAACCTACTCAGAGGAGCTGTGTATACC 555  
 QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyr 89  
 Db 556 CTCATCAATCATCGGTTCTCTGTACAAGGCCATCANGAACTCTGCTGAAGAGTACTAT 615  
 QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109  
 Db 616 GAGNATGTTCCCTGCAAGCTGAGAGACCCAGAGAGTCTTGGGAGGAAGTCTGAGAG 675  
 QY 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128  
 Db 676 TATTCATCTTACATATGCTTCTACAGACCCAGGATGCCGATCCCAAGATGAAG 735  
 QY 129 TyrGluValValPheProHis 135  
 Db 736 TATGAATCTTCTCATGCTCTCAC 756

RESULT 13  
 AY404586  
 LOCUS  
 DEFINITION  
 Homo sapiens GCET2 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 AY404586  
 AY404586.1 GI:39760563  
 GSS.  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 508)  
 Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,  
 Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 508)  
 Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,  
 Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 Location/Qualifiers  
 1..508  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>508  
 /gene="GCET2"  
 /locus\_tag="HCM1940"  
 ORIGIN

```

Qy 120 ProCysSerCys---ThrHieGluHisAspTyrGluValValPheProHis 135
    |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
Db 528 CCCAGGCATGCCGATCCCAAGAGATGAATATGAACTTCTCATGCCTCAC 578

RESULT 10
BM456595          1070 bp      mRNA      linear      EST 05-FEB-2002
LOCUS             AGNCOURT_6409034 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5496261
DEFINITION        5', mRNA sequence.
ACCESSION         BM456595
VERSION           BM456595.1 GI:18505635
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         1 (bases 1 to 1070)
AUTHORS           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL           NIH-MGC http://mgi.nci.nih.gov/
COMMENT           National Institutes of Health, Mammalian Gene Collection (MGC)
                  Unpublished (1999)
                  Contact: Robert Strausberg, Ph.D.
                  Tissue Procurement: Lou Staudt
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM12124 row: h column: 22
                  High quality sequence stop: 648.
FEATURES          Location/Qualifiers
                    1..1070
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5496261"
                        /tissue_type="lymphoma, cell line"
                        /lab_host="DH10B (phage-resistant)"
                        /clone_lib="NIH MGC 85"
                        /notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
                        Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
                        Average insert size 1.867 kb. Library enriched for
                        full-length clones and constructed by Life Technologies.
                        Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores: 9.76e-09 Length: 1070
Pred. No.: 166.50 Matches: 48
Score: 52.23% Conservative: 34
Percent Similarity: 30.57% Mismatches: 52
Best Local Similarity: 23.29% Indels: 23
Query Match: 4 Gaps: 6
DB:

US-10-005-907-2 (1-135) x BM456595 (1-1070)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 136 ATGGGAATTCCTGCTGAGAGAAACAGTTTCAGGCGGCGAGCAACACTCAAGAGATG 195

Qy 19 Pro-----LysLysGlyAsnProAspGluArgLysArg----- 30
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 196 CCTTGAATGTGAGATGCAAAAGCCCAAGAGAAACATCCAGATCGTGGATCACCAT 255

Qy 31 -----GlnGluMetThrThrPheGluArgLysLeu 40
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 256 ATCCTGAAGGCTTTCTGCTCCATGTAAGAAATACTATTITGAAAGAGG--- 312

Qy 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 313 CAAGATTCCTCCCAAGAAATGAAGAATGTCTACTCTCCATCCAGGACAAATGTTGAC 372

```

```

Qy 61 GlySerGlySerGluGluValCysTyrThrValLeuAsnHis---IleProHisGlnArg 79
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 373 CAGACCTACTCAGAGGAGCTGTCTATACCTCATCATCGGGTTCTCTGTACAAGG 432

Qy 80 SerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 433 CCATCAGGGAACCTCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGACTGAGAGACCC 492

Qy 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 493 AGAGAGTCTCTGGGAGGAAGTGAAGTGAATGATTCACCTTCATACATATGCTTCTACAGAC 552

Qy 120 ProCysSerCys---ThrHieGluHisAspTyrGluValValPheProHis 135
    |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
Db 553 CCCAGGCATGCCGATCCCAAGAGATGAATATGAACTTCTCATGCCTCAC 603

BM455198          1193 bp      mRNA      linear      EST 05-FEB-2002
LOCUS             AGNCOURT_6405645 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500141
DEFINITION        5', mRNA sequence.
ACCESSION         BM455198
VERSION           BM455198.1 GI:18504238
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         1 (bases 1 to 1193)
AUTHORS           Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE             Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL           NIH-MGC http://mgi.nci.nih.gov/
COMMENT           National Institutes of Health, Mammalian Gene Collection (MGC)
                  Unpublished (1999)
                  Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Lou Staudt
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM12134 row: j column: 14
                  High quality sequence stop: 592.
FEATURES          Location/Qualifiers
                    1..1193
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5500141"
                        /tissue_type="lymphoma, cell line"
                        /lab_host="DH10B (phage-resistant)"
                        /clone_lib="NIH MGC 85"
                        /notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
                        Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
                        Average insert size 1.867 kb. Library enriched for
                        full-length clones and constructed by Life Technologies.
                        Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores: 1.12e-08 Length: 1193
Pred. No.: 166.50 Matches: 48
Score: 52.23% Conservative: 34
Percent Similarity: 30.57% Mismatches: 52
Best Local Similarity: 23.29% Indels: 23
Query Match: 4 Gaps: 6
DB:

US-10-005-907-2 (1-135) x BM455198 (1-1193)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 65 ATGGGAATTCCTGCTGAGAGAAACAGTTTCAGGCGGCGAGCAACACTCAAGAGATG 124

```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue  
 Genoscope.  
 Direct Submission  
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

FEATURES  
 source  
 1..1566  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DG007YD15"  
 /tissue type="B cells (Ramos cell line)"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

Alignment Scores:  
 Pred. No.: 7.12e-09 Length: 1566  
 Score: 169.50 Matches: 47  
 Percent Similarity: 51.61% Conservative: 33  
 Best Local Similarity: 30.32% Mismatches: 54  
 Query Match: 23.71% Indels: 21  
 DB: 3 Gaps: 5

US-10-005-907-2 (1-135) x CR625820 (1-1566)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluGluAsnGlnLysLysPro--- 19  
 Db 115 ATGGGAATTCCTGCTGAGAGAAACAGCGGCGAGACACTCAAGAGATCGCTTG 174  
 Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 175 AATGTGAGATGCAAGACCCCAACAGACAGACATCCAGATGCTGGGATCCATATCGCT 234  
 Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
 Db 235 GAAGGGTGTTCGCTTCCATCGGAAAAAATACTCATTTTGAAGAAGG---CAAGAT 291  
 Qy 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluArgLys 62  
 Db 292 TCCCAACAGAAATGAAAGATGTCTACTCTCCATCCCATCCAGACAAATGTTGACCGACC 351  
 Qy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81  
 Db 352 TACTCAGAGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGTACAGGCCATCA 411  
 Qy 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 101  
 Db 412 GGGAACTGCTGAGAGTACTATGAGAAATGTTCCCTGCAAGCTGAGAGACCCAGAG 471  
 Qy 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 121  
 Db 472 TCCTTGGGAGGAACTGAGACTGAGTATGATTCATCTTACATATGCTTCTACAGACCC 531  
 Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 532 CATGCCCGATCCCAAGAGATGAATATGAATTCCTCATGCCCTCAC 576

## RESULT 9

BM904106 1000 bp mRNA linear EST 12-MAR-2002  
 LOCUS  
 DEFINITION  
 AGENCOURT.6692745 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:5499556  
 5', mRNA sequence.  
 ACCESSION  
 BM904106  
 VERSION  
 BM904106.1 GI:19354351  
 KEYWORDS  
 EST.

## SOURCE

ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1000)  
 NIH-MGC http://mgs.nci.nih.gov/  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12133 row: b column: 05  
 High quality sequence stop: 698.

## FEATURES

## source

1..1000  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5499556"  
 /tissue type="lymphoma, cell line"  
 /lab host="DH10B (phage-resistant)"  
 /clone lib="NIH\_MGC\_85"  
 /notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.867 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 8.95e-09 Length: 1000  
 Score: 166.50 Matches: 48  
 Percent Similarity: 52.23% Conservative: 34  
 Best Local Similarity: 30.57% Mismatches: 52  
 Query Match: 23.29% Indels: 23  
 DB: 4 Gaps: 6

US-10-005-907-2 (1-135) x BM904106 (1-1000)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18  
 Db 111 ATGGGAATTCCTGCTGAGAGAAACAGTTTCAGCGCGCAGCAGAACACTCAAGAGATG 170  
 Qy 19 Pro-----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 171 CCTTGAATGTGAGATGCAAGACCCCAACAGACACATCCAGATGCTGGATCACCAT 230  
 Qy 31 -----GlnGluMetThrThrPheGluArgLysLeu 40  
 Db 231 ATCGCTGAAGGGTGTTCGCTTCCATCGGAAAAAATACTCATTTTGAAGAAGG--- 287  
 Qy 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlu 60  
 Db 288 CAAGATTCCTCAAAACGAAATGAAAGATGTCTACTCTCCATCATCGGTTCTGTGACAGG 347  
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArg 79  
 Db 348 CAGACCTACTCAGAGGAGCTGTGTATACCTCATCATCATCGGTTCTGTGACAGG 407  
 Qy 80 SerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99  
 Db 408 CCATCAGGAACTCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGAGCTGAGAGACC 467  
 Qy 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119  
 Db 468 AGAGAGTCTCTGGGAGGAACTGAGACTGAGTATTCACTTCTACATATGCCTTCTACAGAC 527

ORIGIN		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4342719"	
		/tissue_type="lymphoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_85"	
		/note="Organ: lymph; vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
US-10-005-907-2 (1-135) x BG106563 (1-1095)			
QY	1	MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluGluAsnGlnLysLysPro	19
Db	43	ATGGGAAATTCCTCTGAGAGAAACACGGCGGACGAGACACTCAAGAGATGCCTTGG	102
QY	20	-----LysLysGlyAsnProAspGluGluArgLysArg-----	30
Db	103	AATGTGAGAAATGCAAGACCCCAACACAGAGAACATCCAGATGCTGGGATCACCATATCGCT	162
QY	31	-----GlnGluMetThrThrPheGluArgLysLeuGlnAsp	42
Db	163	GAAGGGTGTTCCTGCTCCATGGAATAAAATACTATTTCGAAAGAGG---CAAGAT	219
QY	43	GlnAspLysLysSerClnGluValSerThrSerAsnGlnGluAsnGlnLysSer	62
Db	220	TCCCAAAACGAAATGAAAGATGCTATCTCTCCATCCAGACAAATGTTGACCAAGACC	279
QY	63	GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer	81
Db	280	TACTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTCTCTGTACAAGGCCATCA	339
QY	82	LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln	101
Db	340	GGGAACCTCTGCTGACAGACTACTATGAGAAATGTTCCCTGCAAGACTGAGAGACCCAGAG	399
QY	102	PheArgGluArgSerGlnGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys	121
Db	400	TCCTTGGGAGGAACTGAGACTGAGTATTCATCTTCATATGCTTCTACAGACCCGAGG	459
QY	122	SerCys---ThrHisGluHisAspTyrGluValValPheProHis	135
Db	460	CATGCCCGATCCCGAGAGATGAATATGAACTTCTCATGCTCTCAC	504
RESULT 7			
BF795688		1124 bp mRNA linear EST 12-JAN-2001	
LOCUS		602259560F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342719 5',	
DEFINITION		mRNA sequence.	
ACCESSION		BF795688.1 GI:12100659	
VERSION		EST.	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 1124)	
AUTHORS		NIH-MGC http://mgc.nci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strauberg, Ph.D. Email: cspbs-remail.nih.gov Tissue Procurement: Louis Staudt, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9958 row: h column: 16 High quality sequence stop: 715. Location/Qualifiers	
FEATURES		1..1124	
source			
Alignment Scores:			
Pred. No.:	4,648-09	Length:	1124
Score:	169.50	Matches:	47
Percent Similarity:	51.61%	Conservative:	33
Best Local Similarity:	30.32%	Mismatches:	54
Query Match:	23.71%	Indels:	21
DB:	2	Gaps:	5
US-10-005-907-2 (1-135) x BF795688 (1-1124)			
QY	1	MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluGluAsnGlnLysLysPro	19
Db	118	ATGGGAAATTCCTCTGAGAGAAACACGGCGGACGAGACACTCAAGAGATGCCTTGG	177
QY	20	-----LysLysGlyAsnProAspGluGluArgLysArg-----	30
Db	178	AATGTGAGAAATGCAAGACCCCAACACAGAGAACATCCAGATGCTGGGATCACCATATCGCT	237
QY	31	-----GlnGluMetThrThrPheGluArgLysLeuGlnAsp	42
Db	238	GAAGGGTGTTCCTGCTCCATGGAATAAAATACTATTTCGAAAGAGG---CAAGAT	294
QY	43	GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer	62
Db	295	TCCCAAAACGAAATGAAAGATGCTATCTCTCCATCCAGACAAATGTTGACCAAGACC	354
QY	63	GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer	81
Db	355	TACTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTCTCTGTACAAGGCCATCA	414
QY	82	LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln	101
Db	415	GGGAACCTCTGCTGACAGACTACTATGAGAAATGTTCCCTGCAAGACTGAGAGACCCAGAG	474
QY	102	PheArgGluArgSerGlnGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys	121
Db	475	TCCTTGGGAGGAACTGAGACTGAGTATTCATCTTCATATGCTTCTACAGACCCGAGG	534
QY	122	SerCys---ThrHisGluHisAspTyrGluValValPheProHis	135
Db	535	CATGCCCGATCCCGAGAGATGAATATGAACTTCTCATGCTCTCAC	579
RESULT 8			
BF795688		1566 bp mRNA linear HTC 21-JUL-2004	
LOCUS		CR625820 full-length cDNA clone CSODG007YD15 of B cells (Ramos cell line) of	
DEFINITION		Homo sapiens (human).	
ACCESSION		CR625820	
VERSION		CR625820.1 GI:50506627	
KEYWORDS		HTC; CNSLT cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		1 (bases 1 to 1566)	
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
JOURNAL		Full-length cDNA libraries and normalization	
REMARK		Unpublished	
Contact		Feng Liang Email : fliang@lifetech.com URL :	



US-10-005-907-2 (1-135) x BM479887 (1-1039)

VERSION AL560376.3 GI:46185747  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 917)

AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31284507.

CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 1964.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DG007CB080P1&c=1964.r.

FEATURES  
 source Location/Qualifiers

1..917

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DG007YD15"

/tissue\_type="B CELLS (RAMOS CELL LINE)"

/cell\_line="RAMOS CELL LINE"

/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Alignment Scores:  
 Pred. No.: 3 57e-09 Length: 917  
 Score: 169.50 Matches: 47  
 Percent Similarity: 51.61% Conservative: 33  
 Best Local Similarity: 30.32% Mismatches: 54  
 Query Match: 23.71% Indels: 21  
 DB: 1 Gaps: 5

US-10-005-907-2 (1-135) x AL560376 (1-917)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19  
 Db 115 ATGGGAAATTCCTGCTGAGAGAAACAGCGCGCAGAGAACTCAAGAGATGCCTTG 174  
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 175 AATGTGAGATGCAAGAGCCCAACAGAGAACTCCAGATGCTGGGATCACCATATCGCT 234  
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
 Db 235 GAAGGGTGTTCCTGCCTTCCATCGAAAGAAATACTCATTTTGAAGAGAGG---CAAGAT 291  
 QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer 62  
 Db 292 TCCCAAAACGAAATGAAGATGTCTACTTCTCCATCCCAAGGACATGTTGACGAGACC 351  
 QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81  
 Db 352 TACTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTCTCTGTACAGGCCATCA 411  
 QY 82 LeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101  
 Db 412 GGGAACTCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGAGCTGAGAGACGAGAG 471

QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121  
 Db 472 TCCITGGGAGGACTGAGACTGAGTATTCTCATATGCTTCTACAGACCCAGG 531  
 QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 532 CATGCCCGATCCCGAGAGATGAATGATGAATCTTCTATGCCTTCAC 576

RESULT 4  
 BG389415  
 LOCUS  
 DEFINITION

602414106F1 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:4522512 5',

mRNA sequence.

ACCESSION BG389415

VERSION BG389415.1 GI:13282861

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 924)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Place: LLNL0423 row: 1 column: 01

High quality sequence stop: 726.

FEATURES

source Location/Qualifiers

1..924

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4522512"

/tissue\_type="embryonal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 92"

/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_MGC Library."

Alignment Scores:  
 Pred. No.: 3 6e-09 Length: 924  
 Score: 169.50 Matches: 47  
 Percent Similarity: 51.61% Conservative: 33  
 Best Local Similarity: 30.32% Mismatches: 54  
 Query Match: 23.71% Indels: 21  
 DB: 4 Gaps: 5

US-10-005-907-2 (1-135) x BG389415 (1-924)

ORIGIN

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19  
 Db 96 ATGGGAAATTCCTGCTGAGAGAAACAGCGCGCAGAGAACTCAAGAGATGCCTTG 155  
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 156 AATGTGAGATGCAAGAGCCCAACAGAGAACTCCAGATGCTGGGATCACCATATCGCT 215  
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
 Db 216 GAAGGGTGTTCCTGCCTTCCATCGAAAGAAATACTCATTTTGAAGAGAGG---CAAGAT 272

/tissue type="from acute myelogenous leukemia"  
 /lab host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_55"  
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);  
 Site 1: SfiI (ggcgctctggcc); Site 2: SfiI  
 (ggcattatggcc); Double-stranded cDNA was prepared from  
 cell line RNA. 5' and 3' adaptors were used in cloning as  
 follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and  
 3' adaptor sequence:  
 5'-ATTCTAGAGCGGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo  
 Alto, CA)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.78e-73 Length: 527  
 Score: 715.00 Matches: 135  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-005-907-2 (1-135) x BF242113 (1-527)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysProLys 20  
 Db 45 ATGGGAAATTTCTCTCGGAAATCTAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 104  
 QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
 Db 105 AAAGGAAACCCAGATGAGAAAGAAACCGCAGGAAATGACTATTGAAAGAAACTT 164  
 QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60  
 Db 165 CAAGATCAAGATAAGAAAGCCAAAGAGTTTCATCCACTTCTATCAGGAAACCGAGAT 224  
 QY 61 GlySerGlySerGluGluValCysTyrThrValLeuAsnHisIleProHisGlnArgSer 80  
 Db 225 GGCAGTGGTCTGGAAGAGTGTGTACACTGTCTAATTAATCACATCCCCCATCAGATCC 284  
 QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
 Db 285 TCCCTGAGCTCCATGATGATGGCTATGAGAACATTCACCTCCCAAGAAAGTGAGA 344  
 QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
 Db 345 CAGTTTAGAGAAAGGTCAGAGACAGATAATGCCCTTCTTAGGACTTCTGTAGTGGCCT 404  
 QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 405 TGTTCTGCACCCATGAGCATGATTATGAAGTTGTGTGTTCCACAC 449

## RESULT 2

BF797507 1045 bp mRNA linear EST 12-JAN-2001  
 LOCUS 602257174F1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4340323 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF797507  
 VERSION BF797507.1 GI:12102561  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

## REFERENCE

1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (bases 1 to 1045)  
 NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 cDNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHAM9952 row: d column: 20  
 High quality sequence stop: 687.  
 Location/Qualifiers  
 1..1045  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4340323"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_85"  
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.867 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

FEATURES  
source

## Alignment Scores:

Pred. No.: 1.44e-09 Length: 1045  
 Score: 173.50 Matches: 48  
 Percent Similarity: 51.61% Conservative: 32  
 Best Local Similarity: 30.97% Mismatches: 54  
 Query Match: 24.27% Indels: 21  
 DB: 2 Gaps: 5

US-10-005-907-2 (1-135) x BF797507 (1-1045)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysPro--- 19  
 Db 103 ATGGGAAATTTCTCTGCTGAGAGAAACAGCGCGCAGAGAACACTCAAGAGATGCCTGG 162  
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 Db 163 AATGTGAGATGCAAGACCCCAACAGAGAACATCCAGATGCTGGGATCCATATCGCT 222  
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
 Db 223 GAAGGGTGTTCCTGCTTCCATCGAAAAAATACTCATTTTTCGAAAGAGG---CAAGAT 279  
 QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAenGluAenGlySer 62  
 Db 280 TCCCAAAACAAATAAGAAAGATGTCTATCTCTCCCATCCAGGACAAATGTTGACCAGACC 339  
 QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81  
 Db 340 TACTCAGAGGAGCTGTGCTATACCTTCATCAATCATCTGGGTTCCTGTCAAGGCCATCA 399  
 QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101  
 Db 400 GGGAACTCTCTGAAGAGTACTATGAGATGTTCCCTGCAAGAGCTGAGAGACCAGAGAG 459  
 QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121  
 Db 460 TCCTGGGAGGAGCTGAGAGCTGAGTATTCTTCTACATATGCCTTCTACAGACCCAGG 519  
 QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 Db 520 CATGCCCGATCCCGAGAGATGAATGATGAATTCATGCTTCATGCCTCAC 564

## RESULT 3

AL560376 917 bp mRNA linear EST 02-APR-2004  
 LOCUS AL560376 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA  
 DEFINITION clone GS0DG007YD15 5-PRIME, mRNA sequence.  
 ACCESSION AL560376

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 19, 2004, 01:11:39 ; Search time 2705 Seconds  
(without alignments)  
1818.619 Million cell updates/sec

Title: US-10-005-907-2  
Perfect score: 715  
Sequence: 1 MGNVLLRLKSLCIGENQKPK.....SVSRPCSTHEHDYEVVFP 135

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cpn2.1/uspto.spool/US10005907/runat.16112004.140457.6176/app.query.fasta.1.327  
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -TRANSHUMAN40.cdi -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10005907 @CGN.1.1.3437 @runat.16112004.140457.6176 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic.\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	527	2	BF242113
2	173.5	24.3	1045	2	BF797507
3	169.5	23.7	917	1	AL560376
4	169.5	23.7	924	4	EG389415
5	169.5	23.7	1039	4	EG389415
6	169.5	23.7	1095	4	EG106563
7	169.5	23.7	1124	2	BF795688
8	169.5	23.7	1566	3	CR625820
9	166.5	23.3	1000	4	EN904106

10	166.5	23.3	1070	4	BM456595
11	166.5	23.3	1193	4	BM455198
12	159.5	22.3	1085	4	BM455407
13	156.5	21.9	508	9	AY404586
14	156.5	21.9	683	2	AW963366
15	156.5	21.9	876	5	BQ214406
16	150.5	21.0	1041	4	BQ24891
17	149	20.8	1165	4	BM458339
18	146.5	20.5	1059	4	BM459235
19	145	20.3	495	1	AA311038
20	142	19.9	526	6	CB286394
21	133.5	18.7	522	4	BG149093
22	128	17.9	825	6	CB320849
23	127	17.8	520	4	BG145241
24	127	17.8	524	4	BG276802
25	127	17.8	526	4	BG276546
26	125	17.5	661	2	BB629482
27	125	17.5	3317	3	AK036553
28	122	17.1	560	5	BU696543
29	118.5	16.6	815	5	BU932945
30	118	16.5	828	5	BX390865
31	116.5	16.3	998	4	BG176880
32	112	15.7	479	5	BU429114
33	110.5	15.5	539	4	BG276071
34	109.5	15.3	538	2	AW519547
35	108	15.1	1015	4	BM477520
36	105.5	14.8	1527	9	AG032114
37	104.5	14.6	502	9	AY404588
38	104	14.5	591	2	BF797042
39	103.5	14.5	927	4	BG177483
40	102.5	14.3	833	7	CK397921
41	102.5	14.3	965	5	BX391573
42	101.5	14.2	288	2	AW480906
43	101.5	14.2	875	9	AG181241
44	100.5	14.1	508	9	CE686854
45	100	14.0	594	8	BH085248

## ALIGNMENTS

RESULT 1

BF242113

LOCUS

DEFINITION

BF242113

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF242113 527 bp mRNA linear EST 14-NOV-2000  
601880401F1 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:4109064 5',  
RNA sequence.  
BF242113 GI:11156040  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 527)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue procurement: ATCC  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LINC993 row: a column: 01  
High quality sequence stop: 518.  
Location/Qualifiers  
1. 527  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4109064"



[illegible]









QY 45 LysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluSerGlySer 64



CC encode. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, medical imaging, identification of mutations responsible for genetic disorders or other traits, assessing biodiversity and producing many other types of data and products dependent on DNA and amino acid sequences. They are also useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell disorders (e.g. osteoporosis, osteoarthritis, bone degenerative disorders, periodontal disease, liver fibrosis, infections (e.g. viral, fungal or bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis). Sequences ABX04986-ABX05511 represent human polynucleotides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

XX SQ Sequence 775 BP; 226 A; 202 C; 160 G; 187 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4, 7e-08 Length: 775  
Score: 149.00 Matches: 37  
Percent Similarity: 60.1% Conservatives: 28  
Best Local Similarity: 34.2% Mismatches: 39  
Query Match: 20.8% Indels: 4  
DB: 8 Gaps: 4

US-10-005-907-2 (1-135) x ABX05108 (1-775)

QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50  
DB 329 AAAAAAATACTCATTTTGAAGAAGAGG---CAAGATTCCCAAAACGAAATGAAGAATG 385

QY 51 SerSerThrSer---AsnGlnGluAsnGluAsnGlySerGlySerGluGluValCysTyr 69  
DB 386 TCATCTACTCCATCCAGCAGGACAAATGTTGACAGACCTACTCAGAGGAGCTGTCTAT 445

QY 70 ThrValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerSerAsnAspAspGly 88  
DB 446 ACCCTCATCATCATCGGGTTCTCTGTACAGGCCCATCGGGAACCTCTGCTGAAGAGTAC 505

QY 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108  
DB 506 TATGAGAAATGTTCCCTGCAAGCTGAGACCCAGAGAGTCTTGGGAGGAGTACTGAGACT 565

QY 109 GluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHis 127  
DB 566 GAGTATTCACTTCTACATATGCTTCTACAGACCCCGAGGATGCCCGATCCCGAAGAT 625

QY 128 AspTyrGluValValPheProHis 135  
DB 626 GAATATGAACCTTCTCATGCTCTAC 649

RESULT 6  
AAZ24890  
ID AAZ24890 standard; DNA; 1892 BP.  
XX AC  
XX AAZ24890;  
XX DT  
XX 02-DEC-1999 (first entry)  
XX DE Human secreted protein gene 80 clone HLMW54.  
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy; diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia; developmental abnormality; foetal deficiency; blood; allergy; renal; ds; immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS; cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus; osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion; endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX OS Homo sapiens.  
XX PN W09947540-A1.

QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
DB 1156 AATGTGAGAAATGCAAAAGCCCAACAGAGAAATCCAGATGCTGGATCACCATATCGCT 1215

QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
DB 1216 GAAGGGTGTCTGCTCCATGGAAGAAATATCTCATTTTGAAGAAGAGG---CAAGAT 1272

QY 43 GlnAspLysSerGlnGluValSerThrSer---AsnGlnGluAsnGluAsnGly 61  
DB 1273 TCCCAAAACGAAATGAAAGATGTCATCTACTCCATCCAGCAGGACAAATGTTGACGAG 1332

QY 62 SerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 80  
DB 1333 ACCCTACAGAGGAGCTGTGCTATACCTCATCAATCATCGGGTCTCTGTACAGGCCA 1392

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
DB 1393 TCAGGGAATCTGTGAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCCAGA 1452

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
DB 1453 GAGTCTCTGGGAGGAATGAGACTGAGTATTCATCTTANATATGCTTCTACAGACCC 1512

QY 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
DB 1513 AGGCATGCCGATCCCGAGAGATGAATATGAACCTTTCATGCTCTAC 1560

RESULT 5  
ABX05108  
ID ABX05108 standard; cDNA; 775 BP.  
XX AC  
XX ABX05108;  
XX DT  
XX 17-JAN-2003 (first entry)  
XX DE Human novel polynucleotide #123.

XX KW Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer; neurodegenerative disorder; lymphoid cell disorder; osteoporosis; Parkinson's disease; Alzheimer's disease; bone degenerative disorder; osteoarthritis; periodontal disease; liver fibrosis; viral infection; fungal infection; bacterial infection; autoimmune disease; diabetes; atopic dermatitis.  
XX OS Homo sapiens.  
XX PN WO200274961-A1.  
XX PD 26-SEP-2002.  
XX PF 14-MAR-2002; 2002WO-US005109.  
XX PR 15-MAR-2001; 2001US-00810173.  
XX PA (HYSB-) HYSEQ INC.  
XX PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX DR WPI; 2003-040556/03.  
XX DR P-PSDB; ABU00030.  
XX PT New isolated polypeptides and polynucleotides, useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders, lymphoid cell disorders, bone degenerative disorders, and infections.  
XX FS Claim 1; SEQ ID NO 123; 235pp; English.  
XX The invention relates to human polynucleotides and the polypeptides they

Percent Similarity: 51.61% Conservative: 33  
 Best Local Similarity: 20.32% Mismatches: 54  
 Query Match: 23.71% Indels: 21  
 DB: 8 Gaps: 5

US-10-005-907-2 (1-135) x ACC59951 (1-732)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19  
 134 ATGGGAATCTCTCTGAGAGAAACAGGCGGACAGAACTCAAGAGATGCTTGG 193  
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30  
 194 AATGTGAGAAATCAAGACCCCAACAGAGAAATCCAGATCGTGGATCACCATATCCCT 253  
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42  
 254 GAAGGTTGTTCTGCTTCCATGGGAAAAATACTCATTTTGAAGAGAG--CAAGAT 310  
 QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnSer 62  
 311 TCCCAAAACGAAATGAAGATGATCTACTCCATCCAGGACAAATGTTGACGAGACC 370  
 QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81  
 371 TACTCAGAGGAGCTGTGTATACCTCATCATCGGTTCTCTGACAGGCCATCA 430  
 QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101  
 431 GGAAGACTCTGTGAGAGTACTATGAGAATGTTCCCTCAAAGCTGAGAGACCCAGAG 490  
 QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121  
 491 TCCTTGGAGGAAGTGGAGCTGAGTATTCATCTTACATATGCTTCTACAGACCCAGG 550  
 QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135  
 551 CATGCCGATCCCCAGAGATGATATGAACTTCTCATGCTCTCAC 595

#### RESULT 4

ABL99791

ID ABL99791 standard; cDNA; 1643 BP.

AC ABL99791;

DT 03-OCT-2002 (first entry)

XX Human secretory polynucleotide (sptm) 46.

XX Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;

XX SPTM-related disease; somatic gene therapy; germline gene therapy;

XX severe combined immunodeficiency; intracellular parasite protection;

XX fungal parasite; protozoan parasite; cell proliferative disorder; cancer;

XX immune disorder; AIDS; neurological disorder; Parkinson's disease;

XX motor neuron disorder; demyelinating disease; multiple sclerosis;

XX meningitis; abscess; prion diseases; cerebral palsy;

XX neuroskeletal disorder; peripheral nervous system disorder;

XX dermatomyositis; polymyositis; myopathy; myasthenia gravis;

XX mental disorder; Tourette's syndrome.

XX Homo sapiens.

OS WO200220756-A2.

XX 14-MAR-2002.

XX 30-AUG-2001; 2001WO-US027297.

XX 05-SEP-2000; 2000US-0229747P.

XX 05-SEP-2000; 2000US-0229748P.

XX 05-SEP-2000; 2000US-0229749P.

XX 05-SEP-2000; 2000US-0229750P.

XX 05-SEP-2000; 2000US-0229751P.

PR 05-SEP-2000; 2000US-0230016P.  
 PR 05-SEP-2000; 2000US-0230583P.  
 PR 06-SEP-2000; 2000US-0230505P.  
 PR 06-SEP-2000; 2000US-0230514P.  
 PR 06-SEP-2000; 2000US-0230515P.  
 PR 06-SEP-2000; 2000US-0230517P.  
 PR 06-SEP-2000; 2000US-0230518P.  
 PR 06-SEP-2000; 2000US-0230519P.  
 PR 06-SEP-2000; 2000US-0230596P.  
 PR 06-SEP-2000; 2000US-0230597P.  
 PR 06-SEP-2000; 2000US-0230599P.  
 PR 06-SEP-2000; 2000US-0230610P.  
 PR 06-SEP-2000; 2000US-0230864P.  
 PR 06-SEP-2000; 2000US-0230865P.  
 PR 06-SEP-2000; 2000US-0230988P.  
 PR 06-SEP-2000; 2000US-0230989P.  
 PR 06-SEP-2000; 2000US-0230990P.  
 PR 07-SEP-2000; 2000US-0230896P.  
 PR 07-SEP-2000; 2000US-0230897P.  
 PR 07-SEP-2000; 2000US-0230951P.  
 PR 07-SEP-2000; 2000US-0231163P.  
 PR 07-SEP-2000; 2000US-0231832P.

(INCY-) INCYTE GENOMICS INC.

Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;

Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;

Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;

Gerstein EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;

Marwaha R, Chen AU, Chang SC, Au AP, Inman RR;

WPI; 2002-315658/35.

P-PSDB; ABB97794.

Polynucleotide sequences encoding human secretory proteins useful for

gene therapy of e.g. genetic deficiency disorders, cancers, and diseases

caused by intracellular parasites.

Claim 1; Page 275-276; 585pp; English.

The invention comprises the amino acid and coding sequences of human

secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are

useful for treating a disease or condition associated with the expression

of functional SPTM. The SPTM DNA sequences are useful for somatic or

germline gene therapy to correct a genetic deficiency (e.g. severe

combined immunodeficiency). The SPTM DNA sequences are also useful in

providing protection against intracellular parasites (e.g. fungal

parasites and protozoan parasites). The SPTM DNA and protein sequences

are also useful for diagnosing cell proliferative disorders, cancer,

immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's

disease), motor neuron disorders, demyelinating diseases (e.g. multiple

sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,

neuroskeletal disorders, peripheral nervous system disorders,

dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental

disorders (e.g. Tourette's syndrome). cDNA sequences ABL99746 - ABL99929

represent human secretory polynucleotides of the invention

Sequence 1643 BP; 413 A; 382 C; 384 G; 460 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 8.15e-09 Length: 1643

Score: 159.00 Matches: 47

Percent Similarity: 50.64% Conservative: 32

Best Local Similarity: 30.13% Mismatches: 55

Query Match: 22.24% Indels: 22

DB: 6 Gaps: 6

US-10-005-907-2 (1-135) x ABL99791 (1-1643)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19

DB 1096 ATGGGAATCTCTCTGAGAGAAACAGGCGGACAGAACTCAAGAGATGCTTGG 1155

CC response in a patient. The protein can also serve as a target that  
 CC modulate gene expression or activity and as an antigen to raise  
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying  
 CC agents that modulate expression of the protein or agents, such as  
 CC agonists or antagonists. The agonists or antagonists are useful for  
 CC modulating biological activity and function of (II) and thus are useful  
 CC for alleviating disease conditions such as allergic hypersensitivity,  
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis  
 XX  
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.5e-75 Length: 3762  
 Score: 715.00 Matches: 135  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x ABN81319 (1-3762)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20  
 DB 25 ATGGGAAATTATCTCTCGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAACGCCAAG 84  
 QY 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
 DB 85 AAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACATT 144  
 QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60  
 DB 145 CAAGATCAAGATGAAGAAAGCCAGAGTTTCTATCCACTTCTAATCAGGAAACGAGAAAT 204  
 QY 61 GlySerGlySerGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80  
 DB 205 GGCAGTGGTTCGAGAGGTGCTACACTGTCAATTAATCACATCCCCATCAGAGATCC 264  
 QY 81 SerLeuSerSerAsnAspLysGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
 DB 265 TCCTGAGCTCCAATGATGCTGATGAGAACATTCCTCTCAAGAGAAAGTGAGA 324  
 QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
 DB 325 CAGTTTAGAGAAAGGTGAGACAGAGATATGCCCTCTTAGGACTTCTGTAGAGCCT 384  
 QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135  
 DB 385 TGTTCTCGACCATGAGCATGATTATGAAGTTGTGTTCACAC 429

RESULT 3  
 ACC59951  
 ID ACC59951 standard; cDNA; 732 BP.

AC ACC59951;  
 XX  
 DT 07-JUL-2003 (first entry)  
 XX

Human IRAP-28 encoding cDNA SEQ ID NO:63.

XX Human; immune response associated protein; IRAP; anti-HIV; anti-allergic;  
 KW anti-inflammatory; anti-naemic; anti-parkinsonian; neurotropic; antithyroid;  
 KW anticonvulsant; antiarteriosclerotic; antiasthmatic; immunosuppressive;  
 KW cytotatic; hepatotropic; dermatologic; anti-diabetic; nephrotropic;  
 KW antitumor; thymomimetic; neuroprotective; osteopathic; antiarthritic;  
 KW antiparasitic; antihelminthic; antiparasitic; uropathic; protozoacide;  
 KW antirheumatic; haemostatic; antibacterial; virucide; ophthalmological;  
 KW fungicide; gene therapy; immune system disorder; neurological disorder;  
 KW developmental disorder; muscle disorder; cell proliferative disorder;  
 KW infection; gene; ss.

XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH

CDS 134..670  
 /\*tag= a  
 /product= "IRAP-28"  
 WO2003025542-A2.  
 PD 27-MAR-2003.  
 PF 19-SEP-2002; 2002WO-US029979.  
 XX 21-SEP-2001; 2001US-0324034P.  
 PR 05-OCT-2001; 2001US-0327395P.  
 PR 12-OCT-2001; 2001US-0328923P.  
 PR 19-OCT-2001; 2001US-0342810P.  
 PR 09-NOV-2001; 2001US-0344468P.  
 PR 21-NOV-2001; 2001US-0332140P.  
 PR 07-DEC-2001; 2001US-0340282P.  
 PR 09-JAN-2002; 2002US-0347693P.  
 PR 20-FEB-2002; 2002US-0358279P.  
 PR 01-MAR-2002; 2002US-0361088P.  
 PR 15-MAR-2002; 2002US-0364494P.  
 PR 10-MAY-2002; 2002US-0379876P.  
 PR 11-JUN-2002; 2002US-0388180P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX Ho A, Baughn MR, Becha SD, Burford N, Elliott VS, Emerling BM;  
 Forsythe IU, Gorvad AE, Griffin JA, Hafalia AJA, Honchell CD;  
 Ison CH, Burdill JD, Blake JJ, Lal PG, Lee EA, Marquis JP;  
 Lehr-Mason PM, Lee S, Sprague WW, Swarnakar A, Tang YT, Tran B;  
 Tran UK, Bhatia U, Chawla NK, Warren BA, Zheng W, Xu Y, Yue H;  
 WPI: 2003-363161/34.  
 P-PSDB; ABR43232.

New human immune response associated proteins and polynucleotides, useful  
 for diagnosing, treating or preventing immune system disorders, e.g. AIDS  
 or anemia, cell proliferative disorders, e.g. cancer, or neurological  
 disorders.  
 Claim 5; Page 208; 213pp; English.

ACC59924 to ACC59958 encode the human immune response associated proteins  
 given in ABR43205 to ABR43239, designated IRAP-1 to IRAP-29 (I). (I) have  
 anti-HIV, anti-allergic, anti-inflammatory, antinaemic, antiparkinsonian,  
 neurotropic, anticonvulsant, antiarteriosclerotic, antiasthmatic, antitumor,  
 immunosuppressive, antithyroid, cytotatic, hepatotropic, dermatologic,  
 antidiabetic, nephrotropic, thymomimetic, neuroprotective, osteopathic,  
 antirheumatic, antiparasitic, antihelminthic, antiparasitic, virucide,  
 ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,  
 protozoacide and fungicide activities, and can be used in gene therapy.  
 Human IRAP polynucleotides, agonists and antagonists are useful for  
 diagnosing, treating or preventing disorders associated with aberrant  
 expression of IRAP, such as immune system disorders (e.g. AIDS, asthma,  
 allergies, autoimmune thyroiditis, contact dermatitis, Crohn's disease,  
 diabetes mellitus, glomerulonephritis, Good pasture's syndrome, gout,  
 Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,  
 multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's  
 syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis),  
 neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or  
 epilepsy), developmental disorders (e.g. renal tubular acidosis, anaemia  
 or mental retardation), muscle disorder (e.g. cardiomyopathy, myocarditis  
 or dermatomyositis), cell proliferative disorders (e.g. arteriosclerosis,  
 atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal  
 haemoglobinuria, polycythaemia vera, psoriasis, primary  
 thrombocytopenia or cancer), or viral, bacterial, fungal, parasitic,  
 protozoan or helminthic infections

SQ Sequence 732 BP; 213 A; 178 C; 178 G; 163 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.55e-10 Length: 732  
 Score: 169.50 Matches: 47



XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Meng G;  
PI Haley-Vicente D, Dmanac RT;  
XX WPI; 2003-371981/35.  
DR P-PSDB; ABC31731.  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX Claim 1; SEQ ID NO 842; 1185pp; English.  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human cDNA sequence of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9 55e-76 Length: 667  
Score: 715.00 Matches: 135  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-005-907-2 (1-135) x ADC30760 (1-667)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysProLys 20  
Db 186 ATGGGAATTTATCTCTGGGAAACTCAGTTGCTGGGAGAGATCAAGAAGACCCCAAG 245  
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
Db 246 AAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTGAAAGAAACCTT 305  
QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAenGluAen 60  
Db 306 CAGATCAAGATAGAAACCCAGGAGTTTCACCTCTTAATCAGGAAACGAGAAT 365  
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80

Db 366 GGCAGTGTGTTCTGAAGAAGTGTGTACACTGTCTAATATACATCCCATCCAGATCC 425  
QY 81 SerLeuSerSerAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
Db 426 TCCCTGAGCTCCAATGATGATGGCTATGAGAACATTGACTCCCTCAAGGAAAGTGA 485  
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120  
Db 486 CAGTTTAGAGAAAGGTTCAGAGACAGATATGCCCTCTTAGGACTTCTGTAGTAGGCCT 545  
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135  
Db 546 TGTTCCTGCACCCCATGAGCATGATTATGAAGTTGTGTTCACAC 590

RESULT 2  
ABN81319  
ID ABN81319 standard; cDNA; 3762 BP.  
XX AC ABN81319;  
XX 30-AUG-2002 (first entry)  
DE Human mast cell related gene MC1 SEQ ID NO 1.  
KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;  
KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;  
KW gene; ss.  
OS Homo sapiens.  
XX PH Key Location/Qualifiers  
FT CDS 25..432  
FT /\*tag= a  
FT /product= "MC1"  
XX WO200246389-A2.  
XX 13-JUN-2002.  
XX 07-DEC-2001; 2001WO-US046180.  
XX 08-DEC-2000; 2000US-0251835P.  
XX 14-MAR-2001; 2001US-0275479P.  
XX 28-MAR-2001; 2001US-0279115P.  
XX 02-APR-2001; 2001US-0280143P.  
XX (UNIO ) UCB SA.  
XX Nocka K, Pirozzi G, Einstein R;  
XX WPI; 2002-508560/54.  
XX P-PSDB; ABB77569.  
XX Novel isolated nucleic acids that are differentially expressed in mast  
PT cells in patients with allergic hypersensitivity, encoding proteins  
PT associated with mast cell regranulation and allergic hypersensitivity.  
XX Claim 1; Page 95-97; 119pp; English.  
XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),  
CC corresponding to genes differentially expressed in mast cells following  
CC activation or in patients with allergic hypersensitivity disease, (I)  
CC that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of  
CC (II) if at least 6 amino acids. (II) is useful for identifying binding  
CC partners. (I) or (II) is useful for diagnosing or treating a disease  
CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,  
CC urticaria or atopic dermatitis or mastocytosis) in a subject which  
CC involves determining the level of expression of (I) or (II). A computer  
CC system, comprising a database containing information identifying the  
CC expression level in a tissue or at least one mast cell of (I), is useful  
CC for presenting information to identify the relative expression level of  
CC (I). (II) is used as a marker to detect, diagnose or identify an allergic

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 18, 2004, 21:26:34 ; Search time 424 Seconds  
(without alignments)  
1671.396 Million cell updates/sec

Title: US-10-005-907-2  
Perfect score: 715  
Sequence: 1 MGNVLLRLKSLGENQKPK.....SVSRPCSTHEHDYEVVFP 135

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO spool/US10005907/runat\_16112004\_140457.6157/app\_query.fasta\_1.327  
-DB=N Geneseq 23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10005907 @CGN\_1.1.470 @runat\_16112004\_140457.6157 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	667	10	ADC30760 Human nov
2	715	100.0	3762	6	ABN81319 Human mas
3	169.5	23.7	732	8	ACC59951 Human IRA
4	159	22.2	1643	6	AD199791 Human sec
5	149	20.8	775	8	ABX05108 Human nov
6	98	13.7	1892	2	AZ224890 Human sec

7	98	13.7	1892	8	ADA39940	Human sec
8	98	13.7	1892	10	ADC73578	Human sec
9	98	13.7	1892	10	ADD37623	Human sec
10	98	13.7	1892	10	ADA56130	Gene enco
11	97	13.6	3957	12	ADQ24942	Human sof
12	96.5	13.5	584	10	ADE62022	Rat gene
13	96	13.4	2301	2	AAV20445	Human c-t
14	96	13.4	2301	3	AAO09303	Human Trk
15	93	13.0	1824	4	ABL27495	Drosophil
16	92	12.9	654	5	AA571153	DNA enco
17	92	12.9	654	5	AA575467	DNA enco
18	92	12.9	654	5	AA569553	DNA enco
19	91.5	12.8	2862	12	ADO00884	Human hom
20	91.5	12.8	3182	5	ABX71302	Human kid
21	91.5	12.8	3186	12	ADQ24514	Human sof
22	91.5	12.8	5109	10	ADF81540	Leukaemia
23	91.5	12.8	5261	5	ADL45687	Human ova
24	91.5	12.8	7488	6	ABL33112	Human imm
25	90.5	12.7	651	12	ADJ67693	Kifc DNA
26	90	12.6	1671	4	AAH99858	Human pro
27	90	12.6	1899	2	AA506785	Human adu
28	90	12.6	1912	3	AA250927	Human pro
29	90	12.6	7992	3	ABK32895	DNA enco
30	90	12.6	8083	5	AA570692	DNA enco
31	90	12.6	8257	9	ACA90154	Human cdn
32	90	12.6	8257	12	ADQ17931	Human sof
33	90	12.6	8491	10	ADI02719	Human cdn
34	90	12.6	8493	8	ACA03930	cDNA down
35	90	12.6	8503	4	AA158253	Human pol
36	90	12.6	8503	5	ADQ98460	DNA enco
37	90	12.6	8503	9	ADB48220	Novel hum
38	90	12.6	8517	12	ADQ22543	Human sof
39	90	12.6	8527	2	AA160039	Human pol
40	89.5	12.5	1027	4	AAK4690	CBCALD05
41	89.5	12.5	1027	6	ABK47580	cDNA enco
42	89.5	12.5	1071	4	ABA83301	Human sec
43	89.5	12.5	1071	9	ACH04802	Novel hum
44	89.5	12.5	1071	9	ACD44612	Human cdn
45	89.5	12.5	3719	10	ADI40373	Human pur

ALIGNMENTS

RESULT 1  
ADC30760  
ID ADC30760 standard; cDNA; 667 BP.  
XX  
AC ADC30760;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human novel cDNA sequence, SEQ ID NO:842.  
XX  
KW Human; diagnostic; drug screening; forensics; gene mapping;  
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KW ulcers; osteoporosis; autoimmune disease; cancer;  
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;  
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KW gene therapy; chromosome 1; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO2003029271-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 24-SEP-2002; 2002WO-US030474.  
XX  
PR 24-SEP-2001; 2001US-0324631P.  
XX  
PA (HYSE-) HYSEQ INC.



LOCATION: 1, 2, 3, 4, 5, 6, 5250, 5251, 5252, 5253, 5254, 5255, 5256,  
LOCATION: 5257, 5258, 5259, 5260, 5261  
OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-19577

Alignment Scores:  
Pred. No.: 0.702 Length: 5261  
Score: 91.50 Matches: 32  
Percent Similarity: 47.90% Conservative: 25  
Best Local Similarity: 26.89% Mismatches: 45  
Query Match: 12.80% Indels: 17  
DB: 10 Gaps: 4

US-10-005-907-2 (1-135) x US-09-814-353-19577 (1-5261)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20  
DB 610 ATCGGCCAGTCGTTGTAAGAAGAACAGACCTAACCGGAGGAAACGAGCTGCTGGAG 669

QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
DB 670 GAG-----CAGGTGAACACATCAGGAGGAGGTGTCTCAGCTCCGGCATGAGCTG 720

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60  
DB 721 TCCATGAAGATGAGCTGCTTCACTTACACAGCGCTCGGAGGAG-----768

QY 61 GlySerGlySerGluGluValCysTyrThrValLeuAsnHisLeuProHisGlnArgSer 80  
DB 769 ---AGTGAGCCCGAGTCCGTTTCTCAACC-----CCGTTGAAGAGGAAT 810

QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArg 100  
DB 811 GAGTCGCTCTCTCAGTCAGTCCAGAAATTTTCATTGGATTCTCTCAAGAAGAGCTGAA 870

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119  
DB 871 GACCTTGAAGAG-----GAGATGTTGTACTTCGATCCGAGGCCACCCAG 915

RESULT 13  
US-10-311-455-1085/c  
Sequence 1085, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Cytosine Methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
PRIOR FILING DATE: 2002-12-16  
PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1085  
LENGTH: 7488  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1085  
Alignment Scores:  
Pred. No.: 1.15 Length: 7488  
Score: 91.50 Matches: 29  
Percent Similarity: 44.22% Conservative: 36  
Best Local Similarity: 19.73% Mismatches: 51  
Query Match: 12.80% Indels: 31

DB: 15 Gaps: 5  
US-10-005-907-2 (1-135) x US-10-311-455-1085 (1-7488)

QY 14 GluAsnGlnLysLys-----Pro 19  
DB 6080 GAAATACTATAAAAAATACGACTCACTTCTTCCGCTCGCAAAAATACACTACTATAA 6021

QY 20 LysLysGlyAsnProAspGluGluArgLysArgGlnGlu-----MetThrThr 35  
DB 6020 AAACAAATAACCCCGATAAAAAATAAATAAACACACCTTAACCTCTCTACTATAAC 5961

QY 36 PheGluArg-----LysLeuGlnAspGlnAspLysLysSer 47  
DB 5960 TTTAAAAATCTAACCCAAAATCTAAAAACCTTTAAAAATACGAAAAAATCTTAAAAATCT 5901

QY 48 GlnGluValSerSerThrSerAsnGlnGluAsnGluAsnGlySerGlySerGluGluVal 67  
DB 5900 AAAAACCCCTTACCAACGAATAACCTTAAAAATAACCAAAAAATAACCTTACCTCGCCCT 5841

QY 68 CysTyrThr---ValLeuAsnHisLeuProHisGlnArgSerSerLeuSerSerAsnAsp 86  
DB 5840 CTAAATTCCTTCTTCTACTCACTATCTCGATACCCACGTAATAAATAAATAAATAAATA 5781

QY 87 AspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106  
DB 5780 CACAATATCTATAAAATTCATAATATTTTAAAAAATCCGTTCTTACTTAAACAAACGAA 5721

QY 107 GluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126  
DB 5720 AAAACGACCCACCACTCGAAAA-----AAACCAATAATAACACACGTTTC 5673

QY 127 HisAspTyrGluValValPhe 133  
DB 5672 CAGATAAAAAAATAATCTAT 5652

RESULT 14  
US-10-261-175A-9  
Sequence 9, Application US/10261175A  
Publication No. US20040038222A1  
GENERAL INFORMATION:  
APPLICANT: DIETRICH, WILLIAM  
APPLICANT: WAITERS, JAMES W.  
TITLE OF INVENTION: ANTHRAX SUSCEPTIBILITY GENE  
FILE REFERENCE: 56491(71250)  
CURRENT APPLICATION NUMBER: US/10/261,175A  
CURRENT FILING DATE: 2002-09-30  
PRIOR APPLICATION NUMBER: 60/325,864  
PRIOR FILING DATE: 2001-09-29  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 651  
TYPE: DNA  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Kiflc nucleic acid sequence

US-10-261-175A-9  
Alignment Scores:  
Pred. No.: 0.0514 Length: 651  
Score: 90.50 Matches: 39  
Percent Similarity: 36.97% Conservative: 22  
Best Local Similarity: 23.64% Mismatches: 51  
Query Match: 12.66% Indels: 53  
DB: 16 Gaps: 7

US-10-005-907-2 (1-135) x US-10-261-175A-9 (1-651)

QY 16 GlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThr 35  
DB 4 GAGAAACACCCAGCTCGGTGACAGGAGTCCGCGCAAGCGCCGAGACGAGCTGCC 63

```

; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552
; LENGTH: 5109
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1552

Alignment Scores:
Pred. No.: 0.673 Length: 5109
Score: 91.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 15 Gaps: 4

US-10-005-907-2 (1-135) x US-10-172-118-1552 (1-5109)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
Db 610 ATCGGCAGTCGTTGTTGAAGAAGAACAAAGACCCCTAACCGAGAGGAGCTGCTGGAG 669
Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 670 GAG-----CAGGTGGAACACATCAGGAGGAGGTCTCTACGCTCCGCGATGAGCTG 720
Qy 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
Db 721 TCCATGAAGATGAGCTGCTTACAGTCTTACACACGCGCTCGGAGGAG----- 768
Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 769 ---AGTGACCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAAT 810
Qy 81 SerLeuSerSerAsnAspGlyTyrGluAenIleAspSerLeuThrArgLysValArg 100
Db 811 GAGTCGCTCTCTCAGTCAGATTCAGTTCATTGGATTCTCTCAAAAGAACTGAAA 870
Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db 871 GACCTTGAAGAG-----GAGAATGTTGTACTTCGATCCGAGGCCAGCCAG 915

RESULT 11
US-10-342-887-1552
; Sequence 1552, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552
; LENGTH: 5109
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1552

Alignment Scores:
Pred. No.: 0.673 Length: 5109
Score: 91.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 15 Gaps: 4

US-10-005-907-2 (1-135) x US-10-172-118-1552 (1-5109)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
Db 610 ATCGGCAGTCGTTGTTGAAGAAGAACAAAGACCCCTAACCGAGAGGAGCTGCTGGAG 669
Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 670 GAG-----CAGGTGGAACACATCAGGAGGAGGTCTCTACGCTCCGCGATGAGCTG 720
Qy 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
Db 721 TCCATGAAGATGAGCTGCTTACAGTCTTACACACGCGCTCGGAGGAG----- 768
Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 769 ---AGTGACCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAAT 810
Qy 81 SerLeuSerSerAsnAspGlyTyrGluAenIleAspSerLeuThrArgLysValArg 100
Db 811 GAGTCGCTCTCTCAGTCAGATTCAGTTCATTGGATTCTCTCAAAAGAACTGAAA 870
Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db 871 GACCTTGAAGAG-----GAGAATGTTGTACTTCGATCCGAGGCCAGCCAG 915

RESULT 12
US-09-814-353-19577
; Sequence 19577, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19577
; LENGTH: 5261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

```

Db 991 TCATCCGGTCAACACAAATGCTTATGCAAGAGGTAAACAAAGCATTCCTCTCT 1050  
 Qy 79 ArgSerSerLeuSerSerAsnAsp-----AspGlyTyrGluAsnIleAspSerLeuThr 96  
 Db 1051 AGTGATGAAGTAGTAGTATGATGATGAAGGTGAGGAAAGCCCTCTCTAGATGAGCTTCG 1110  
 Qy 97 ArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSer 116  
 Db 1111 GAAGCCGTAGAAATTTTCCAGATGTTTGCCTAAGCAAAAGCTCAACTTAATACTTGG 1170  
 Qy 117 ValSerArgProCysSerCysThrHisGluHisAspTyrGlu 130  
 Db 1171 AAAATAAGTTGGTTAGC-----TCTCAAAATGATTACAA 1206  
 RESULT 8  
 US-10-115-831-54  
 ; Sequence 54, Application US/10115831  
 ; Publication No. US20030219743A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Ren, Feiyan  
 ; APPLICANT: Drmanac, Radoje T.  
 ; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and  
 ; TITLE OF INVENTION: Polypeptides  
 ; FILE REFERENCE: 792CIP2ADIV  
 ; CURRENT APPLICATION NUMBER: US/10/115,831  
 ; CURRENT FILING DATE: 2002-04-02  
 ; PRIOR APPLICATION NUMBER: 09/667,298  
 ; PRIOR FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: 09/577,408  
 ; PRIOR FILING DATE: 2000-05-18  
 ; NUMBER OF SEQ ID NOS: 178  
 ; SOFTWARE: pt\_FL\_genes Version 2.0  
 ; SEQ ID NO 54  
 ; LENGTH: 6324  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(3573)  
 US-10-115-831-54  
 Alignment Scores:  
 Pred. No.: 0.78 Length: 6324  
 Score: 92.00 Matches: 27  
 Percent Similarity: 38.60% Conservative: 17  
 Best Local Similarity: 23.68% Mismatches: 50  
 Query Match: 12.87% Indels: 20  
 DB: 15 Gaps: 3  
 US-10-005-907-2 (1-135) x US-10-115-831-54 (1-6324)  
 Qy 14 GluAsnGlnLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33  
 Db 337 GACAAATGACGAGCCAGGGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396  
 Qy 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluValSerSerThr 53  
 Db 397 GAGGAGATCGAGGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAA 456  
 Qy 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluVal----- 67  
 Db 457 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516  
 Qy 68 -----CysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsn 85  
 Db 517 ATGNAATTGTCACAAAT-----ACTCGAATAATGCAACACACAGAAAGGATGATAACAAT 570  
 Qy 86 AspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal----- 99  
 Db 571 AATGACGAATATGACAAATACGATGAATGTCGCGCAAGTCATTTAAACCTCGGCAAA 630

Qy 100 -----ArgGlnPheArgGluArgSerGluThrGlu 109  
 Db 631 ATCGCTGAGGATGCAGCTACCGGCCAGGAGTGAAGTCAAGAA 672  
 RESULT 9  
 US-10-263-929-22  
 ; Sequence 22, Application US/10263929  
 ; Publication No. US20040067535A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; APPLICANT: Galant, Ron  
 ; TITLE OF INVENTION: Alzheimer's Disease Linked Genes  
 ; FILE REFERENCE: LSD-07417  
 ; CURRENT APPLICATION NUMBER: US/10/263,929  
 ; CURRENT FILING DATE: 2002-10-03  
 ; NUMBER OF SEQ ID NOS: 213  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 22  
 ; LENGTH: 2862  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-263-929-22  
 Alignment Scores:  
 Pred. No.: 0.3 Length: 2862  
 Score: 91.50 Matches: 32  
 Percent Similarity: 47.90% Conservative: 25  
 Best Local Similarity: 26.89% Mismatches: 45  
 Query Match: 12.80% Indels: 17  
 DB: 16 Gaps: 4  
 US-10-005-907-2 (1-135) x US-10-263-929-22 (1-2862)  
 Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20  
 Db 394 ATCGGCCAGTCTGTTGAAGAAGAACAAAGACCCCTAACCCAGAGGAGGAGGAGGAGGAG 453  
 Qy 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40  
 Db 454 GAG-----CAGGTGGAACACATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504  
 Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60  
 Db 505 TCCATGAGGATGAGCTGCTTCACTTCTACACAGCGCTCGGAGGAG----- 552  
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80  
 Db 553 ---AGTGAGCCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAAT 594  
 Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100  
 Db 595 GAGTCGCTCTCTCTCAGTCCAGAAATTAATTCATTTCATTTGATTCCTTCAAGAGAGCTGAA 654  
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119  
 Db 655 GACCTTGAAGAG-----GAGAAATGTTGTACTTCGATCCGAGGCCAGCCAG 699  
 RESULT 10  
 US-10-172-118-1552  
 ; Sequence 1552, Application US/10172118  
 ; Publication No. US20030224374A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dai, Hongyue  
 ; APPLICANT: He, Yudong  
 ; APPLICANT: Linsley, Peter  
 ; APPLICANT: Mao, Mao  
 ; APPLICANT: Roberts, Chris  
 ; APPLICANT: Van 't Veer, Laura  
 ; APPLICANT: Van de Vijver, Marc  
 ; APPLICANT: Bernards, Rene  
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
 ; FILE REFERENCE: 9301-175-999





```

; PRIOR FILING DATE: 1998-03-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-653-595-90

Alignment Scores:
Pred. No.: 0.0237 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 16 Gaps: 4

US-10-005-907-2 (1-135) x US-10-653-595-90 (1-1892)
QY 5 LeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnPro 24
DB 572 CTCCTACAAATCCTCTGCTGGAGCCAGGAGCAAGCCAGAGCAGCAGCAGAGGAG 631
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
DB 632 CAAGAGGTGGACACAGGAGGAGCGGACACAAAGAACACACAGCAGGAAGGGGCGAGAA 691
QY 45 LysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGluAsnGlySerGlySer 64
DB 692 CAGGAAGAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 752 AAGGAGGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
QY 81 ---SerLeuSerSerAsnAspAspGlyTyr-----ArgSerGluThrGluTyrAlaLeuLeuArg 89
DB 812 GAATCTCTATCTTCTAAACCTTCTCTTTGCTCCCGGTACGAGAGTAGAGTCTACT 871
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
DB 872 CCTATGATAATGAGAGAACATCCAGGAGCTCATTCGATCAGCCAGGAAATAGATGAATG 931
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
DB 932 AATGAAATATATGATGAGAACTCTACTGGAGAAACCAAAACCTCGCAGCTCTGCAG 991
QY 115 ThrSerValSerArgProCysSerCys 123
DB 992 CTGCCCCACACAGAGCTTGCTGTGTC 1018

RESULT 5
US-10-242-943-3
; Sequence 3, Application US/10242943
; Publication No. US20030087412A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, MeiJa
; TITLE OF INVENTION: Nlk1 PROTEIN AND Nlk1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 Nlk1 protein complexes
; CURRENT APPLICATION NUMBER: US/10/242,943
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US/09/167,206
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: November 19, 2004, 02:51:09 ; Search time 1525 Seconds  
(without alignments)  
478.354 Million cell updates/sec

Title: US-10-005-907-2  
Perfect score: 715  
Sequence: 1 MGNVLLRLKSLGENQKPKP.....SVSRPCSTHEDHYEVVFP 135

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3627888 seqs, 2701811610 residues  
Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_model/US10005907/runat\_16112004\_140459\_6250/app\_query.fasta\_1.327  
-DB=published\_applications NA -QFWT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITIS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-NAXLEN=2000000000 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	3762	15	US-10-005-907-1
2	98	13.7	1173	18	US-10-425-115-147529
3	98	13.7	1892	10	US-09-397-945-90
4	98	13.7	1892	16	US-10-653-595-90
5	96	13.4	2301	14	US-10-242-943-3
6	93	13.0	1903	16	US-10-424-599-8569
7	92	12.9	1941	18	US-10-425-115-147532
8	92	12.9	6324	15	US-10-115-831-54
9	91.5	12.8	2862	16	US-10-263-529-22
10	91.5	12.8	5109	15	US-10-172-118-1552
11	91.5	12.8	5109	16	US-10-342-887-1552
12	91.5	12.8	5261	10	US-09-814-353-19577
13	91.5	12.8	7488	15	US-10-311-455-1085
14	90.5	12.7	651	16	US-10-261-175A-9
15	90.5	12.7	2693	17	US-10-437-963-64985
16	90.5	12.6	1173	13	US-10-027-632-10372
17	90	12.6	1173	15	US-10-027-632-10372
18	90	12.6	1671	16	US-10-296-115-693
19	90	12.6	1899	10	US-09-746-783-77
20	90	12.6	7992	10	US-09-893-519A-140
21	90	12.6	8491	15	US-10-133-013-260
22	90	12.6	8493	13	US-10-071-766-51
23	90	12.6	8503	15	US-10-037-270-130
24	90	12.6	8503	15	US-10-117-722-130
25	89.5	12.5	1071	10	US-09-933-767-118
26	89.5	12.5	1071	14	US-10-004-860-118
27	89.5	12.5	1071	14	US-10-023-282-118
28	89.5	12.5	1107	16	US-10-424-599-112896
29	89.5	12.5	9415	15	US-10-311-455-268
30	89	12.4	864	15	US-10-320-800-49
31	89	12.4	3522	18	US-10-425-115-19622
32	89	12.4	5262	18	US-10-425-115-19617
33	88.5	12.4	3370	17	US-10-437-963-83303
34	88.5	12.4	3966	17	US-10-437-963-83303
35	88	12.3	858	17	US-10-425-115-147531
36	88	12.3	1647	18	US-10-425-115-147531
37	88	12.3	9934	15	US-10-087-805-576
38	87	12.2	3720	13	US-10-076-622-576
39	87	12.2	3720	15	US-10-124-805-576
40	87	12.2	3720	15	US-10-437-963-41655
41	87	12.2	3873	17	US-10-451-467A-659
42	87	12.2	81440	17	US-10-175-523-71
43	87	12.2	123192	14	US-10-424-599-23224
44	86.5	12.1	776	16	US-10-415-134-3
45	86.5	12.1	1181	16	US-10-415-134-3

ALIGNMENTS

RESULT 1  
US-10-005-907-1  
; Sequence 1, Application US/10005907  
; Publication No. US20030166881A1  
; GENERAL INFORMATION:  
; APPLICANT: Union Chimique Belge, S.A.  
; APPLICANT: Pirozzi, Gregory  
; APPLICANT: Einstein, Richard  
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST  
; TITLE OF INVENTION: ACTIVATION  
; FILE REFERENCE: 053529-5005  
; CURRENT APPLICATION NUMBER: US/10/005.907  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 3762  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: